普通高等教育"十五"国家级规划教材配套用书

数学建模竞赛

—浙江大学学生获奖论文点评（1999—2004）

杨启帆 何 勇 谈之奕 编

浙江大学出版社



ISBN 7-308-04369-X /O · 327 定价: 28.00元

普通高等教育"十五"国家级规划教材配套用书

数学建模竞赛

——浙江大学学生获奖论文点评

（1999—2004）

杨启帆　何　勇　谈之奕　编

浙江大學出版社

图书在版编目（CIP）数据

数学建模竞赛:浙江大学学生获奖论文点评／杨启帆 何勇 谈之奕编. —杭州:浙江大学出版社，2005.5
ISBN 7-308-04369-X

Ⅰ.数... Ⅱ.杨... Ⅲ.数学模型－文集
Ⅳ.022-53

中国版本图书馆 CIP 数据核字（2005）第 084058 号

责任编辑 李桂云
出版发行 浙江大学出版社
（杭州浙大路 38 号 邮政编码 310027）
（E-mail: zupress@mail.hz.zj.cn）
（网址:http://www.zjupress.com）
排　　版 杭州好友排版工作室
印　　刷 杭新印务有限公司
开　　本 880mm×1230mm 1/32
印　　张 19
字　　数 564 千
版 印 次 2005 年 7 月第 1 版 2005 年 7 月第 1 次印刷
书　　号 ISBN 7-308-04369-X/O·327
定　　价 28.00 元

前 言

浙江大学从1983年起开设数学建模课，最初几年，该课程为数学系选修课和必修课。当时我们开设这门课程的目的主要有两个，一是想让学生通过该课程的学习和参加一定的数学建模实践，亲身体会到数学虽是一门依赖于抽象思维的学科，但决不是一座空中楼阁，其研究的问题大多具有很强的实际背景和广泛的应用前景。脱离生产实际，数学就会失去生命力，失去推动其发展的原动力，从而引起他们对实际课题研究的兴趣。其二，鉴于当时有部分数学系学生认为学习数学比较乏味，学了也不知道有什么用处的现象，希望通过这门课程的开设使这部分学生了解到打好数学基础的重要性，并激发起他们学习和应用数学知识的积极性。数学建模课的开设受到了学生们的普遍欢迎，基本上达到了预期的目的，学生学习兴趣非常浓厚并积极参加建模实践，有的还写出了课题研究报告或研究论文。

在数学系开课的成功使我们受到了很大的启发与鼓舞，也使我们认识到高等教育除了传授知识以外，还应当注重对学生综合素质的培养，尤其应当创造一定的机会和环境，让学生们去运用书本知识，在运用过程中开拓他们的进取精神、创新精神和竞争意识，而数学建模课的开设恰好为他们提供了这样的机会。此后，我们很快把在数学系开课搞试点所取得的经验推广到全校，在浙江大学开出了面向不同对象的各种数学建模课程，形成了一定的规模。目前，我们每年开设的数学建模课有数学系必修课，竺可桢学院混合班（我校尖子学生班，每届约有200人左右）、工程高级班、理科基地班必修课，工科研究生学位课等，每年听课学生有1000人左右。近两年来，我们又在两个本科二级学院中开出了旨在培养应用型人才的数学建模学位课，同样取得了提高学生综合素质、激发他们的创新

能力和竞争意识的良好作用。此外,我们还在校内开设了面向全校学生的选修课,不定期地举办一些讲座,通过这些方法,吸引更多学生参加数学建模学习,投身数学建模实践活动。

1992年开始,教育部高教司和中国工业与应用数学学会共同举办了每年一次的全国大学生数学建模竞赛。我们借此良机,将数学建模教学、组织学生参加建模实践和组织指导学生参加国内外大学生建模竞赛有机结合起来,努力营造我校数学建模系列教学的良好环境。在不断改革教学内容与教学方法的同时,我们又每年举办了全国竞赛和国际竞赛培训班各一期。学生参加建模实践和建模竞赛的积极性很高,自发组织起建模兴趣小组,找题目参与课外课题研究、定期举办研讨班。同学们积极报名参加在每年5月份学校科技活动月期间举行的全校数学建模竞赛(竞赛期间学生照常上课,7～10天后提交研究论文)。暑假期间同学们自发地参加华东地区高校数学建模联赛,并在竞赛中取得了相当不错的成绩。同学们还踊跃报名参加国内外大学生数学建模竞赛,并在竞赛中发挥出色,取得突出成绩。在1999年和2003年的美国大学生数学建模竞赛中,我校学生两度获得特等奖兼INFORMS奖,即美国运筹与管理学会奖(注:到目前为止,亚洲地区高校仅获过这两项INFORMS奖)。以1999—2004年的6年为例,除获2项国际竞赛INFORMS奖外,我校学生还获得了国际竞赛一等奖17项,二等奖9项;全国一等奖21项、全国二等奖26项。获奖级别之高、数量之多均居国内高校前列。对我校学生在国内外数学建模竞赛中所取得的优异成绩,人民日报、光明日报、文汇报、中国教育报、中央人民广播电台、省市报刊电台、香港文汇报及纽约时报等国内外新闻机构均多次作过报道。我们承担的"数学建模教学与实践"教改项目也于2001年通过国家级鉴定,并获得了国家级教学成果二等奖。2003年,我校数学建模课程经同行专家评审,被教育部授予国家级精品课程。现在,数学建模教学、学生建模实践和参赛活动已在浙江大学蓬蓬勃勃地开展起来,成为我校本科教学中的一个亮点,被学校确定为我校惟一的本科生"十五"规划重点建设项目。

学生们在建模实践中提高了综合素质、增长了才干,在此过程中,我们也积累起了大批学生获奖论文,逐步丰富了我们的教学内容。学生的

获奖论文是同学们辛勤劳动的结晶，也是后届学生学习建模时可以参考的范文，为了更好地发挥这批资料的作用，我们特从获奖论文中挑选出部分汇编成册，并加以适当点评，予以正式出版。我们编著本书的主要目的是想为参加国内外大学生数学建模竞赛的同学或对这些课题的研究感兴趣的读者提供一本较好的参考资料。我们认为，指导教师虽然有可以根据经验把握住课题研究的方向，常常也能区分出论文的优劣。但由于时间与精力的限制，指导教师一般不可能像参赛学生那样去深入地开展研究，从而也很难写出优于获奖论文的文章来。因此，虽然获奖论文并非竞赛题的标准答案(事实上，作为数学建模竞赛题，也不应该有所谓的标准答案)，每一篇论文中仍可能有这样那样的错误或不足之处，但它们仍可以作为学习数学建模的范文，其中必有初学者可以借鉴和学习之处。由于篇幅的限制，我们不可能汇编全部获奖论文。例如，2000 年和 2001 年，我校学生在参加美国大学生数学建模竞赛时，全部参赛队均获得了国际竞赛一等奖(注：当时规定每一学校只能选派 6 个参赛队参加竞赛)，如将这些论文全部汇编进去是不现实的。

为了使本书具有更大的参考价值，我们在编著时遵循了以下几条原则：

(1)每一竞赛题最多选用一篇获奖论文；

(2)每一篇被选用的论文都是获得全国一等奖或国际竞赛一等奖以上的论文；

(3)对原文不作修改，以保持文章的原汁原味。

全书共分两大部分，第一部分为 1999—2004 年间的全国大学生数学建模竞赛优秀论文，共 11 篇，第二部分为 1999—2004 年间美国大学生数学建模竞赛获国际一等奖以上的论文，共 13 篇。

本书是我们正在编著的“十五”国家级规划教材和国家级精品课程教材《数学建模》一书的配套教材之一(注：另一配套教材为《数学建模实验》，也将于近期出版)，既可用作数学建模教师教学参考之用，也可用于数学建模竞赛培训或用作工程技术人员的参考书籍。

近年来，以提高学生综合素质与创新能力为目的的大学生数学建模竞赛已成为国内外规模最大的大学生学科竞赛，可以说涉及面最广、影响

面最大，希望本书的出版能有利于这些活动的开展。

我校对学生开展数学建模教学虽然起步较早，也取得了一定的成绩，但与其他数学课程相比还显得不够成熟，尚有许多地方需要不断改进。数学建模教学难度较大，教学内容涉及面广而且更新快，教学方法灵活多变，没有固定的模式。教师既要在课堂教育中传授知识与技能，达到示范、启发以至抛砖引玉的作用，又要激励学生参加建模实践，鼓励他们在建模实践与竞赛中增长真才实学。这种与众不同的教学模式效果虽好，但对教师和学生的要求也比较高，需要我们不断地去探索教学改革的新思路、新途径。我们同样希望能通过本书的出版，促进我们与国内外同行之间的交流。兄弟院校的老师、同学在数学建模教学与参赛方面都有许多成功的经验值得我们好好学习，我们热切期望能得到来自兄弟院校的宝贵意见。我们的看法与做法只是我们的一家之言，缺点与错误在所难免，不妥之处敬请读者批评指正。

编　者

2005 年 2 月

于浙江大学求是园

目　录

第一部分　全国大学生数学建模竞赛

第二部分　美国大学生数学建模竞赛

第一部分

全国大学生数学建模竞赛



自动化车床管理

沈　权　肖　菲　王晓初

摘要

本文讨论了在自动化车床管理过程中,知道了零件、刀具等分布规律的前提下,设计刀具更换方案,及用来判断零件是否合格的检查方案,以使得连续生产的工序能达到最好的效益。

文中,我们先由问题提供的刀具故障记录,用分布拟合检验得到刀具寿命的概率分布为以均值为600,方差为196.6的正态分布,建立了两个较合理的模型,较好地描述了原问题,并设计了问题①、②的刀具更换策略,我们给出了两种操作简便的换刀策略,在工程中易于实现并能得到较好的效益。

由模型1,2我们解得工序效益最好的检查间隔分别为:

问题①:检查间隔为19(个零件/件);

问题②:检查间隔为23(个零件/次)。

同时我们对改进工序检验方式给出了两点建议,并给出简单的证明。

文中,又对原问题进行了计算机随机模拟,得到结果与前述理论模型的结果较接近,再一次验证了模型的合理性和正确性。同时通过模拟统计,我们还分析损耗的组成及部分的变化规律,提出了一种简单的近似算法,并且用模拟法对改进工序检查方式给出的建议进行了验证。

本文中采取的一些面向工程的假设与策略,很具有实际意义。

一、问题重述

一道工序用自动化车床连续加工某种零件,由于刀具损坏等原因该



工序会出现故障，其中刀具损坏故障占 95%，其他故障仅占 5%，工序出现故障是完全随机的，假定在生产任一零件时出现故障的机会均相同。工作人员通过检查零件来确定工序是否出现故障。

现积累有 100 次刀具故障纪录，故障出现时该刀具完成的零件数如附表(附表数据略)。现计划在刀具加工一定件数后定期更换新刀具。

已知生产工序的费用参数如下：

故障时产出的零件损失费用 $f=200$ 元/件；

进行检查的费用 $t=10$ 元/次；

发现故障进行调节时恢复正常的平均费用 $d=3000$ 元/次(包括刀具费)；

未发现故障时更换一把新刀具的费用 $k=1000$ 元/次。

(1)假定工序故障时产出的零件均为不合格品，正常时产出的零件均为合格品，试对该工序设计效益最好的检查间隔(生产多少零件检查一次)和刀具更换策略。

(2)如果该工序正常时产出的零件不全是合格品，有 2% 为不合格品；工序故障时产出的零件有 40% 为合格品，60% 为不合格品。工序正常时认为有故障停机产生的损失费用 1500 元/次。对该工序设计效益最好的检查间隔和刀具更换策略。

(3)在(2)的情况下，可否改进检查方式以获得更高的效益。

二、模型的初步分析

(1)从问题给出的刀具故障记录，通过概率论我们可以得到刀具的故障分布函数，通过对数据的观察研究及查阅常用的系统故障分布函数(文[3] p.18)，我们猜想，本问题的刀具的故障分布函数可能为正态分布。设参数为数学期望 μ，方差 σ，用极大似然估计法，得：

$$\hat{\mu}=\frac{1}{100}\sum_{i=1}^{100}X_i=600,\quad \hat{\sigma}=\sqrt{\frac{\sum_{i=1}^{100}(X_i-\hat{\mu})^2}{100}}=195.6$$

应用分布拟合检验的 χ^2 检验法，根据如下定理(文[1]pp.215～216)：



定理:若 n 充分大($n \geqslant 50$),H_0:总体 x 的分布函数为 $F(x)$,则当 H_0 为真时(不论 H_0 中的分布属何种分布),统计量(5.4)总是近似地服从自由度为 $k-r-1$ 的 χ^2 分布,其中,r 是被估计的参数个数。于是,若在假设 H_0 下算得有

$$\chi^2 \geqslant \chi_\alpha^2(k-r-1)$$

其中

$$\chi^2 = \sum_{i=1}^{k} \frac{(f_i - n\hat{p})^2}{n\hat{p}}$$

则在显著水平 α 下拒绝 H_0,否则,就接受 H_0。

我们将刀具故障记录数据分成十组,用 χ^2 检验法,代入 μ,σ 的估计量,得:

$$\chi^2 = 1.7594$$

$$x_{0.05}^2(k-r-1) = \chi_{0.05}^2(10-1-1) = 15.5 > 1.7594$$

所以,将刀具故障分布看成 $X \sim N(600, 195.6^2)$是可以接受的。

(2)进一步对拟合的分布函数进行验证:将题目所给数据用计算机统计,求出其故障的分布概率函数,并将拟合出来的概率分布函数作图比较,如图 1,发现两者吻合的很好,由此更进一步验证了拟合成正态分布的合理性。

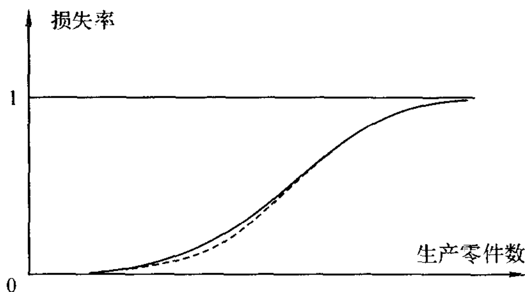


图 1

三、模型假设

(1)刀具损坏故障和其他故障相互独立;

(2)假设单位刀具寿命内,零件抽查的次数足够多,使得可以认为在

任两个相邻抽查之间,刀具故障的概率密度相同;

(3)假设无两种故障(如刀具故障和非刀具故障,两种非刀具故障等)同时产生,也可以认为这种情况是小概率事件;

(4)非刀具故障在一定时期内的概率分布为均匀分布。

四、符号定义

原题中常数的参数意义保留。文中还可能用到的符号有:

w:工序正常而误认为有故障,停机产生的损失费用(元/次);

η_1:工序正常时零件合格的百分比,为常数(=98%);

η_2:工序不正常时零件合格的百分比,为常数(=40%);

E:零件平均工作周期内生产的合格零件数(个);

L:零件平均工作周期内的资金损耗(元);

X_N:工序零件的检查间隔(个/次);

X_M:零件的预防性更换给定间隔(个);

$F(X)$:刀具寿命不超过 x 个零件的概率。

五、模型一

(1)刀具更换策略选定:由于刀具费为1000元/次,比发现故障再调节的费用(3000元/次)要少得多,故我们采用如下一种"预防性刀具更换策略"同"故障后更换"相结合的混合更换策略。

给出一个预防性更换给定间隔 X_M,当处理零件数 $x < X_M$,刀具进行故障后更换,当 $x = X_M$ 且仍正常运行时采取预防性更换,从而使得大批量生产零件时的经济损失最少。

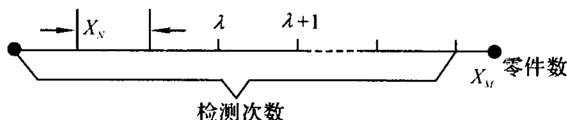


图2 工序说明

(2)由模型假设(4),非刀具故障在一定时期内的概率分布为均匀分布。而且,从总体来看,非刀具故障和刀具故障成一定的比例(1:19)。则



从刀具正态分布规律,可认为非刀具故障在工序进程中分布的概率密度近似如下:

$$f(x)=\frac{1}{19}\cdot\frac{1}{6\sigma}$$

(3)定义 E 为零件平均工作周期内,生产的合格零件数;L 为零件平均工作周期内的资金损耗。则可以将工序效益最好转换为如下规划:

$$\min \frac{L}{E} \tag{1}$$

根据问题可将 L 和 E 表示如下:

$$L = [1-F(X_M)]\cdot\left[K+\left\lfloor\frac{X_M}{X_N}\right\rfloor\cdot t\right] + \sum_{\lambda=0}^{\left\lfloor\frac{X_M}{X_N}\right\rfloor-1}\left[\frac{1}{2}X_N\cdot f+\lambda\cdot t+d+\frac{1}{19}\cdot\frac{1}{6\sigma}(\lambda+1)\cdot X_N\cdot f\right]\cdot [E((\lambda+1)X_N)-F(\lambda X_N)]+L_m$$

$$E = \int_{X_M}^{\infty}X_M\mathrm{d}F + \sum_{\lambda=0}^{\left\lfloor\frac{X_M}{X_N}\right\rfloor-1}[(\lambda+1)X_N-\frac{1}{2}X_N-\frac{1}{19}\cdot\frac{1}{6\sigma}\cdot(\frac{1}{2}X_N)\cdot(\lambda+1)X_N][F((\lambda+1)X_N)-F(\lambda X_N)]+E_m$$

其中,L_m,E_m 表示因取整而造成的截断误差项,其值如下:

$$L_m = \left[F(X_M)-F\left(\left\lfloor\frac{X_M}{X_N}\right\rfloor\cdot X_N\right)\right]\cdot\left[\left\lfloor\frac{X_M}{X_N}\right\rfloor\cdot t + k+\frac{1}{19}\cdot\frac{1}{6\sigma}X_M\cdot f+\frac{\left(X_M-\left\lfloor\frac{X_M}{X_N}\right\rfloor\cdot X_N\right)}{2}\cdot f\right]$$

$$E_m = \left[F(X_M)-F\left(\left\lfloor\frac{X_M}{X_N}\right\rfloor X_N\right)\right]\cdot\left[\left\lfloor\frac{X_M}{X_N}\right\rfloor\cdot X_N + \frac{\left(X_M-\left\lfloor\frac{X_M}{X_N}\right\rfloor X_N\right)}{2}-\frac{1}{19}\cdot\frac{1}{6\sigma}\cdot X_M\cdot\frac{\left(X_M-\left\lfloor\frac{X_M}{X_N}\right\rfloor X_N\right)}{2}\right]$$

(4)由于 L/E 为 X_M 和 X_N 的函数,用 Matlab 编程,对 X_M 和 X_N 进

行网格搜索，求得满足目标函数(1)式的最优解。所得结果如下表：

参数	数值
$\min L/E$(元/个)	4.475
X_M(个)	360
X_N(个)	19
相应的 L(元)	1564.6
相应的 E(个)	350

六、模型二

(1)对于问题的第二问，由于加入了随机的因子，使得问题更为复杂。为了进一步分析方便及减少计算量，我们又进一步作出如下合理假设：

①由于非刀具故障占总故障的百分比很小，仅为5%，因而，我们假设其对于结果影响不大，所以下面我们的模型不考虑非刀具故障。

②在平均的机器周期内，当有两个以上零件不合格时，可以认为此时工序已有故障了。(对此假设的解释和证明见附录)

(2)由上述假设②，我们认为：在工序的进行过程中，一旦发现第二个不合格品时，就认为发现工序故障而调节检修。因非刀具故障为小概率事件，基本可认为是刀具故障，此将作为我们的刀具更换策略之一。

(3)检查到不合格零件时的决策：若是在工序的前半段时间检查到不合格产品，应认为工序不正常的概率不大，不宜认为工序不正常而停机；相反，若是在工序的后半段时间检查到不合格产品，由于时间的积累，工序不正常的概率很大，倾向于停机检查故障，但对于工程实际来说，比如说工人，如何来判断是否需要停机呢？下面给出一个定量判断的方法：

设第 x 个零件被检查为不合格品，则其前工序发生故障的概率(后验概率)可计算如下：

$$F(x)=\frac{F(x)\cdot(1-\eta_2)}{F(x)\cdot(1-\eta_2)+(1-F(x))\cdot(1-\eta_1)}$$

$$=\frac{30}{29F(x)+1}$$

另一方面，若在零件 x 之前，工序已出现故障，则在之后由于零件服



从 $p=\eta_2$ 的 0－1 分布，所以，在 x 以后第一个被检查到的不合格零件与 x 的距离期望为：

$$X_N \cdot \sum_{i=1}^{\infty} \eta_2^{i-1} \cdot (1-\eta_2) \cdot i = \frac{5}{3} \cdot X_N \tag{2}$$

从而由上式得在 x 处决策，将"工序实际有故障而误判为不停机"的损失为

$$\cos t T_1 = \hat{F}(x) \cdot (1-\eta_2) \frac{5}{3} X_N \cdot f$$

进一步可得，在 x 处将"工序实际无故障而误判为停机"的损失为

$$\cos t T_2 = [1-\hat{F}(x)] \cdot w$$

在工序初始阶段，应为 $\cos t T_1 < \cos t T_2$，而零件加工周期快结束时则相反。为了取得最大利益和付出最小损失，当 $\cos t T_1 < \cos t T_2$ 时，应选择不停机；而当 $\cos t T_1 > \cos t T_2$ 时，应选择停机。但作为一般工程人员和一般工人来说，这是一个相当抽象的概念，为此，我们将它转化为感性的、通俗易掌握的标准：给出一阈值 x_0，当被检查到为不合格的零件的序号 $x < x_0$ 时，不停机；当 $x > x_0$ 时，停机。而 x_0 正是当 $\cos t T_1 = \cos t T_2$ 时的零件序号，用 Matlab 编程可得（其中，求 x_0 时要用到的参数 X_N 采用第一小题得到的最优解 $X_N = 19$），得到：对于本问题提供的工序，$x_0 = 165$（个零件）。

（4）刀具更换策略的选定：根据前面的分析，我们选定刀具更换规则如下：

①若刀具工作时间（以零件个数计）到达和超过 X_M 时，换刀。

②若停机并查到刀具出现故障，换刀。

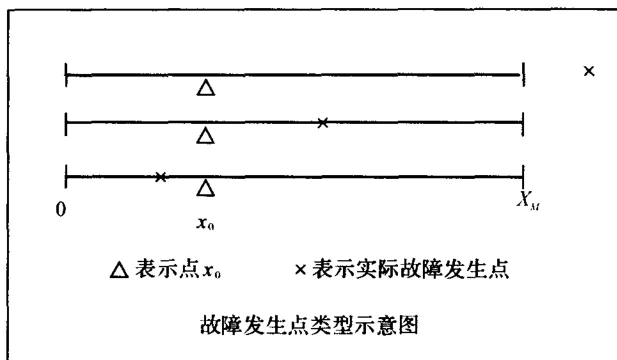
③若到某个零件为止（包括该零件），已查到有两个零件不合格时，则换刀。

（5）与模型一一样，我们也以类似的意义定义 L 和 E，及最优化条件：

$$\min \frac{L}{E} \tag{3}$$

（6）我们按故障发生点的不同来对可能发生的事件进行分类（见图

$\min \frac{L}{E}$



故障发生点类型示意图

图 3　故障发生点类型示意图

3)，对于 L 和 E 的求取，我们在前述假设的条件下，给与适当的简化。如认为 X_N 不是很大，因此在求法中的很多地方，将取整符号去掉了，此时，认为由此引起的截断误差不大。详细如下：

$$L = [1 - F(X_M)]\left[k + \frac{X_M}{X_N}\cdot t + f\cdot(1-\eta_1)\cdot X_M + \eta_1\left(\left\lfloor\frac{X_M}{X_N}\right\rfloor - 1\right)\cdot(1-\eta_1)\cdot\frac{X_M - x_0}{X_N}\cdot w\right] + \sum_{\lambda=0}^{\left\lfloor\frac{X_M}{X_N}\right\rfloor - 1}\{[F((\lambda+1)X_N) - F(\lambda X_N)]\cdot TEMP\}$$

$$X = [1 - F(X_M)]\cdot X_M\cdot\eta_1 + \sum_{\lambda=0}^{\left\lfloor\frac{X_M}{X_N}\right\rfloor - 1}[F((\lambda+1)X_N) - F(\lambda X_N)]\cdot TX$$

其中　$y = (\lambda + 0.5)X_N$

$$TEMP=\begin{cases} d+\eta^{\frac{y_0-x_0}{X_N}-1}\cdot(1-\eta_1)\cdot\dfrac{y_0-x_0}{X_N}\cdot w \\ \quad +(\lambda+0.5)\cdot X_N\cdot(1-\eta_1)\cdot f+TA & (x_0\leqslant y_0) \\ d+(\lambda+0.5)\cdot X_N\cdot(1-\eta_1)\cdot f+TB & (x_0>y_0) \end{cases}$$

$$TX=\begin{cases} (\lambda+0.5)X_N\cdot\eta_1+\left(\dfrac{5}{3}+0.5\right)X_N\cdot\eta_2 & (x_0\leqslant y_0) \\ \left[\dfrac{5}{3}\cdot X_N+x_0-(\lambda+0.5)X_N\right]\cdot\eta_2+(\lambda+0.5)X_N\cdot\eta_1 & (x_0>y_0) \end{cases}$$

$$TA=\begin{cases} \dfrac{5}{3}X_N\cdot(1-\eta_2)\cdot f+\left(\dfrac{5}{3}+0.5+\lambda\right)\cdot t & ((\dfrac{5}{3}+0.5+\lambda)X_N<X_M) \\ [X_M-(0.5+\lambda)\cdot X_N]\cdot(1-\eta_2)\cdot f+\dfrac{X_M}{X_N}\cdot t & ((\dfrac{5}{3}+0.5+\lambda)X_N\geqslant X_M) \end{cases}$$

$$TB=\begin{cases} \left[\dfrac{5}{3}\cdot X_N+x_0-(\lambda+0.5)X_N\right]\cdot(1-\eta_2)\cdot f+\dfrac{\frac{5}{3}X_N+x_0}{X_N}\cdot t & (\dfrac{5}{3}X_N+x_0<X_M) \\ [X_M-(\lambda+0.5)X_N]\cdot(1-\eta_2)\cdot f+\dfrac{X_M}{X_N}\cdot t & (\dfrac{5}{3}X_N+x_0\geqslant X_M) \end{cases}$$

(4)通过 Matlab 用网格搜索法编程得(3)式的解如下表:

参数	数值
minL/E(元/个)	8.74
X_M)(个)	321
X_N(个)	23
相应的 L(元)	2476.7
相应的 E(个)	283

(5)由上面的解可知,与模型一相比,工序正常时的不合格产品率上升;而工序故障时不合格产品率下降。从结果来看,min L/E 增大了将近两倍。$X_M-E=38$ 也比模型一中增大了许多。由此可见,不合格产品中,工序正常时产生的占大部分。η_1 偏离 1 的大小对于结果有着举足轻重的作用。这一点,在下面的计算机模拟中还将得到进一步验证。

七、计算机模型及其对模型的验证

本题为一较复杂的非线性优化问题。为了验证我们所建模型的正确性,我们采用计算机对原问题进行模拟。模拟结果与我们的模型计算结果吻合的相当好,而且通过模拟我们对原问题的性质有了更深入的了解(讨论中将给出原问题的一个简化模型)。

模拟方法及结果:

原问题为一实际的质量管理问题。生产过程可看成一个生产零件的序列,所以我们以零件的序列为模拟序列。生产过程又受到很多随机因素的约束,其中刀具的损害为一主要的随机事件,在以上的分析中我们已经证明了可认为刀具的寿命服从正态分布,所以我们能较容易地标记出刀具损坏的概率密度 f,而其生产第 x 个零件时损坏的概率为 $f(x)$,在模拟程序中我们用一随机函数模拟此事件发生的过程。同样,问题中给出了其他随机事件的概率,我们也可用随机函数来模拟它们发生的过程。对于给定的检查间隔我们随机模拟 2000 次,已能给出较稳定的解答。

问题一的模拟求解:

原问题要求给出一效益最好的检查间隔和刀具更换策略,我们先对检查间隔和刀具更换间隔采用大范围的粗精度搜索找到效益最优解可能存在的范围,再进行较高精度的搜索并求出结果:

换刀间隔	检查间隔	单位零件损耗
360	17	4.52

问题二的模拟求解:

本问题求解受更多的随机事件的影响,如:工序正常时产出有 2% 为不合格品等,我们采用随机函数来模拟这些事件的发生过程,此外我们采用模型二的刀具更换策略,结果如下:

换刀间隔	检查间隔	单位零件损耗
315	21	8.78



通过计算机模拟结果的统计,我们还得到如下一些有意义的结果:

(1)可近似认为工序故障是由刀具故障产生(在模型二中我们作了同样的假设)。

(2)由于在换刀周期内换刀的概率较小,所以可近似认为刀具损坏概率可用 $f(x)$表示(其中 x 为从开始所生产的零件数)。

(3)对于问题二,其损耗由以下这些部分组成:

① $C_1=10$ 元/每次的检查费用;

② $C_2=200$ 元/每个的零件损失费;

③ $C_3=1000$ 元/每次的换刀费;

④ $C_4=1500$ 元/每次的误判费;

⑤ $C_5=3000$ 元/每次的故障恢复费。

通过统计我们得到以上各部分在总费用中的比例,而其随换刀具周期的变化结果如图 4 所示。

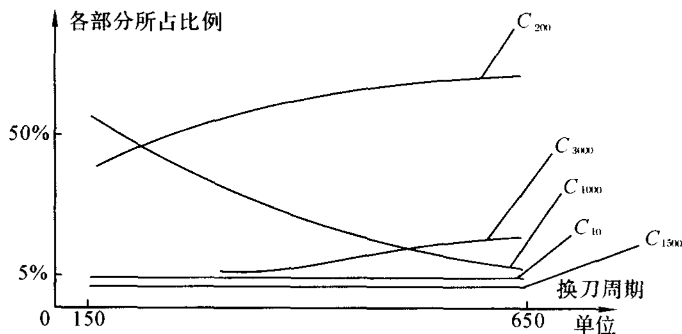


图 4

从图 4 我们可以看到 C_{200}在总的损耗中占很大比例,主要由于在工序正常时也会产生 2% 的次品,如:由于工序发生故障的概率比较小,故可估算得到换刀周期为 300,单位损耗约为:

$$300\times0.02\times200/(300\times0.98)\approx4(\text{元})$$

而总单位费用为 9 元左右,可见生产合格率非常重要。

(4)C_{1500}只占总金额的 3% 左右。

由(1),(2),(4)我们有以下的近似公式:

问题一：

$$L = X_M / X_N \times 10 + p \times 3000 + 1000 + p \times X_N / 2 \times 200$$

$$p = \sum_{i=1}^{X_M} f(i); E = X_M - X_N / 2 \times p$$

以上 X_M 为换刀间隔，X_N 为检查间隔，f 为损耗的概率密度函数。

换刀间隔	检查间隔	单位零件损耗
345	19	4.82

问题二：

$$L = X_M / X_N \times 10 + p \times 3000 + 1000 + p \times X_N \times 5/3 \times 200 \times 0.6 + X_M \times 0.02 \times 200$$

$$p = \sum_{i=1}^{X_M} f(i); E = X_M - X_N / 2 \times p$$

换刀间隔	检查间隔	单位零件损耗
330	15	9.15

由以上的快速算法我们能很快地得到一组较优的解，可见去除一些对结果影响不大的复杂因素，可以达到化繁为简的目的。

八、模型二的改进

(1)不同于前面两个模型，我们认为，检查周期不一定要是一个常数。可以根据故障出现时刻的概率分布来确定。在故障概率大时检查周期短，故障概率小时检查周期长。比如，本文讨论的问题中，由于在小于600之前，刀具故障概率成上升趋势，因而，大体上，检查间隔应是由长逐渐变短的。

(2)由前面文中的讨论可以知道，由于两次检查是被($X_N - 1$)个零件隔开的；因而，实际的不合格产品是检查到的 $1/X_N$。产品中不合格品的比例较大。另一方又由于两个被检查零件间的零件未被检查，对于“接收工序故障信息”有一定的滞后作用，所以，带来很大弊端。



由前面模型二的分析过程中,已经得到了工序不合格时下一个不合格零件与当前零件(设为 x_i)的距离期望为(5/3) * X_N(正如(2)式所示)。为了讨论检查间隔,不妨设 x_i 不合格。

若对于每个零件都检验(相当于 $X_N=1$),其后第一个被观察到的零件距离 x_i 的期望为 5/3。从统计的观点来看,可以认为,只需找两次便可得到第二个不合格的零件,从而换刀进入新的周期。如果这两次检查都合格,则如附录那样可证得 x_i 之前工序故障,故可认为 x_i 前无故障。

综上,我们采取的检查方案为:

在原来的等间距检查的基础上,若检查到某零件 x_i 不合格,则查其下一个 x_{i+1},若 x_{i+1}合格,则再查下一个 x_{i+2},若 x_{i+2}合格,则不再查下一个,认为 x_i 前无故障;否则,若 x_{i+1}和 x_{i+2}中有一个不合格,则换刀进入新的周期。

以下说明检查不合格零件后立即进行下一个零件检查可以减少的损失费:

$$C = \left\{ \sum_{i=0}^{\left\lfloor \frac{X_M}{X_N} \right\rfloor - 2} \left\{ [F((i+1)X_N) - F(iX_N)] \right.\right.$$

$$\times \sum_{j=i+1}^{\left\lfloor \frac{X_M}{X_N} \right\rfloor - 1} (\eta_2 \times (1-\eta_2)^{j-i}) \times f(1-\eta_2)$$

$$\left.\left. \times \frac{5}{3} \times X_N - t \right\} \right. \times \frac{1}{X_N}$$

用 $X_M=321, X_N=23$ 代入计算得:$C=0.15$,此结果与用计算机模拟采用如上策略时,减少的损失费较接近。

九、结果检验和评价

(1)模型的结果分析:

以上我们提出了两个模型并对比计算机模拟检验给出了一近似算法,我们将计算结果归纳如下:

问题一：

	换刀间隔	检查间隔	单位零件损耗
模型一	365	19	4.47
计算机模拟	360	17	4.52
近似算法	345	19	4.82

问题二：

	换刀间隔	检查间隔	单位零件损耗
模型二	321	23	8.74
模型二改进模拟	325	变步长	8.62
计算机模拟	315	21	8.78
近似算法	330	15	9.15

模型的优缺点分析：

①模型的可操作性强，运用我们提出的策略不用作什么计算。

②模型结果较优且和随机模拟的结果吻合得较好。

③我们用计算机模拟不但检验了模型的正确性，而且还分析了损耗组成，提出了一种近似算法，并对模型二的改进提供直观的帮助。

④为了计算方便，对一些非主要因素我们作了近似处理，造成一定的误差。

⑤由于模型考虑的因素较多，所以模型的计算过程较复杂。

⑥虽然通过了假设检验，但认为刀具损耗为正态分布和实际情况仍会存在一定的误差。

十、参考文献

[1]盛骤，谢式千，潘承毅编．概率论与数理统计．北京：高等教育出版社，1989

[2]姜启源编．数学模型．北京：高等教育出版社，1993

[3]周广涛，肖端庄编著．电子设备与系统可靠性设计．《航天工业部第二研究院院刊》编辑部，1984

十一、附录

一个命题的简单证明：

在平均的机器周期内，当有两个以上零件不合格时，可以认为此时工序已有故障了。

证明：

可以用后验概率来说明。

当有两次出现不合格零件(设为第 λ_1 次和第 λ_2 次，$\lambda_1<\lambda_2$)时，在第 λ_2 次之工序有可能故障的概率可计算如下：

$$F'=\frac{[1-F(\lambda_2\cdot X_N)](1-\eta_1)^2}{F(\lambda_1\cdot X_N)(1-\eta_1)^2+[F(\lambda_2\cdot X_N)-F(\lambda_1\cdot X_N)][(1-\eta_2)(1-\eta_1)+F(\lambda_2)](1-\eta_1)^2}$$

$$=\frac{1}{900\cdot\dfrac{E(\lambda_1\cdot X_N)}{1-F(\lambda_2\cdot X_N)}+30\cdot\dfrac{F(\lambda_2\cdot X_N)-F(\lambda_1\cdot X_N)}{1-F(\lambda_2\cdot X_N)}+1}\leqslant\frac{1}{900\cdot F(\lambda_1\cdot X_N)+1}$$

若 $F(\lambda_1X_N)>0.02$，则 $F'<1/(0.02*900+1)=5\%$，可认为是小概率事件，基本不发生。否则，$F(\lambda_1X_N)\leqslant 0.02$，则针对本问题的分布函数，$\lambda_1X_N\leqslant 200$。即前 200 个零件中出现两次不合格。以第一小题优化结果来作为参数，估计其概率。前 200 个零件中，共进行抽查约 $200/X_N\approx 10$ 次，

$$\begin{aligned}F''&=C_{10}^2\cdot[(1-F(x))\cdot(1-\eta_1)+F(x)\cdot(1-\eta_2)]^2\\&\quad\cdot\{1-[(1-F(x))\cdot(1-\eta_1)+F(x)\cdot(1-\eta_2)]\}^8\\&\leqslant 45*[0.02+0.58F(x)]^2\leqslant 45*(0.02+0.58*0.02)^2\\&=0.045<0.05\end{aligned}$$

也属小概率事件，由前述，我们可以认为上述命题成立。

点　评

生产质量管理是概率统计应用较多的一个领域，但长期以来，在我国受重视程度不够，数模赛题中与概率统计相关问题也不多，造成选手们对该项训练有所不足。对此类问题来说，能对数据进行合理的统计分析，给



出符合实际的结论是一项基本功,正确与否直接影响后阶段模型的质量。尽管目前有多种统计分析软件可供选用,但仍需了解基本概念,掌握基本原理,切忌不知所以然地套用现成公式。就本题而言,拟合连续型分布在技术上本不困难,但用题目给出的刀具故障记录拟合所得 $F(x)$ 只是一个条件分布,是在刀具故障、新刀起始点已知条件下的刀具的寿命分布。要得到工序故障间隔(包括刀具故障和非刀具故障)的分布需要对 $F(x)$ 进行修正,这是容易疏忽的。

求解本题两个重要环节是确定效益函数和设计刀具更换策略。本文将工序效益函数表示成“一周期内每生产一合格零件所需的资金”,同时提出了一种把刀具“预防性更换”和“损坏后更换”相结合的“混合更换策略”。论文对问题的求解通过三个途径,一是对理论模型,用计算机数值求值求解非线性规划,二是用计算机模拟的方法验证理论解的合理性,三是设计了计算方便和便于操作的近似算法。论文同时对这三种情况作了比较和讨论。为了求解方便,论文假设刀具故障和非刀具故障不会同时产生,一定程度上增加了误差,稍有不足。

本文获 1999 年全国大学生数学建模竞赛全国一等奖。

钻井布局

沈华品　来　翔　吴以均

摘　要

本文讨论钻井布局模型,采用了**"归一化"**的简化手段,在不同的距离定义下和坐标旋转与否的条件下,给出了钻井布局的最优方案,在计算机上实现了算法,并且用题目给出的数据进行了验证,得出了不同情况下的最优解分别为4和6。对于问题三,我们得到了一般性的判定定理。

一、问题重述

勘探部门在某地区找矿,初步勘探时期已零散地在若干位置上钻井,取得了地质资料。现要在一个区域内按纵横等距的五个点来布置井位,进行"撒网式"全面钻探。

为节省费用,应充分利用旧井,少打新井。

设平面上有 n 个点 P_i,坐标分别为 (a_i,b_i),$i=1,2,\cdots,n$,表示已有的 n 个井位,新布置的井位是一个正方形网格 N 的所有节点,每个格子的边长均为1单位,整个网格在平面上任意移动。若一已知点 P_i 与某个网格节点 X_i 的距离不超过给定误差 $\varepsilon(=0.05$ 单位),则认为该处的旧井资料可利用,不必再打新井。

现欲解决下列问题:

问题一

假定网格的横向和纵向固定,并规定两点间的距离为其横向距离和纵向距离的最大值,在平面上平行移动网格 N,使可利用的旧井数尽可

能达到最大。试提供数值计算方法，并对表一的数值例子用计算机进行计算。

问题二

在欧氏距离的误差意义下，考虑网格的横向和纵向不固定(可以旋转)的情形，给出算法及计算结果。

问题三

如有 n 口旧井，给出判定这些井均可利用的条件和算法(任选一种距离)。

表 1　$n=12$ 个点的坐标值

i	1	2	3	4	5	6	7	8	9	10	11	12
a_i	0.50	1.41	3.00	3.37	3.40	4.72	5.43	5.43	7.57	8.38	8.98	9.50
b_i	2.00	3.50	1.50	3.51	5.50	2.00	6.24	4.10	2.01	4.50	3.41	0.80

二、问题的分析

(1)考虑到钻井的实际情况，利用旧井和钻新井的费用相差很大。故对于本问题，在确定精度和误差范围内，我们应寻求精确最优解和可行方案。

(2)在解决第一个问题的时候，为简化问题，我们采用“归一化”的方法，把正方形网格 N 简化到 $[0,1]\times[0,1]$ 区间的平面范围。取数据的小数部分，作为数据的新坐标，则所有数据落到 $[0,1]\times[0,1]$ 的平面范围。对于这一问题，在“穷举法”的基础上，我们又设计了“边界状态算法”进行二维搜索。

(3)在解决第二个问题的时候，由于坐标系可以旋转，并且转角是连续变化的，所以我们采用“角度离散化”的方法，在 $[0°,90°]$ 范围内等距搜索，以求得近似解。然后，使用“逐步逼近”的方法，逐步减小步长，直到找到符合精度要求的最优解。

(4)对于第三个问题，我们选取了第一种距离定义方式，首先研究了不考虑坐标轴旋转的情况，在“归一化”处理之后，给出了判定两口旧井可以利用的定理一。然后，考虑坐标轴的旋转，给出了判定 n 口旧井均可



利用的定理二,并设计了“组合算法”,具体实现了此算法。

三、基本假设

假设一: 若一已知点 P_i 与某个网格节点 X_i 的距离不超过给定误差 ε(= 0.05 单位),则认为该处的旧井资料可利用,不必再打新井。

假设二: 钻井的实际情况是可以允许一定的误差,钻井定位的结果只需精确到 0.01 单位。

四、符号说明

P_i: 第 i 个点;

(a_i, b_i): 第 i 个点在原坐标系下的坐标;

ε: 给定误差限;

n: 总的点数(井位数)。

五、模型的建立及求解

1. 问题的建模和求解

(1)“归一化”方法

在网格不旋转的情况,进行“归一化”:取数据的小数部分作为数据的新坐标,相当于原网格的各小网格叠合于[0,1]×[0,1]区间的平面范围。此时问题(1)转化为:平移坐标系,使四个小方格内的总点数最大。如图 1 所示。

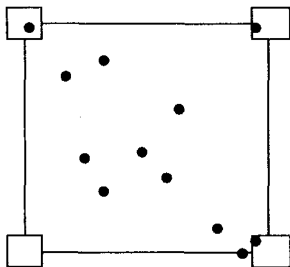


图 1

(2)算法一:穷举法

对于这个问题,考虑到题目中要求的精度为 0.01 单位,所以可以使用“穷举法”进行求解。具体做法是在[0,1]×[0,1]的平面区域内,各点固定,以 0.01 单位为步长,坐标轴沿 X 轴正向和 Y 轴正向平移。每移动一次,记录落在新坐标系下以(0,0),(1,0),(0,1),(1,1)为中心、边长为 2ε 的小正方

形内的个数。最后,各种状态下点数的最大值为最优解。

从上面的算法描述可知,由于精度为 0.01,所以一维方向搜索次数是 100。故总的搜索次数是 $100^2 = 10000$。

(3)算法二:二维边界状态算法

为了简化问题,首先,我们提出了一维边界状态算法,进而得到二维算法。

①一维边界状态算法

以 X 轴为例,仅考虑各点在 X 轴方向上的投影点,固定 X 坐标系及各个投影点的坐标不变,考虑一对长度为 2ϵ 个单位,中心相距为 1 的区间对 E(如图 2)。左区间中心(T_0)初始坐标为 $x_0 = -\epsilon$。利用 E 的正向移动,每次搜寻落在左区间($x_0 - \epsilon, x_0 + \epsilon$)之外且距离该区间右端点最近的一个点,右移区间对 E,使该点刚好在左区间的右端重合。每次移动以后,达到一个新的状态…如此一直进行下去,直到 $x_0 \geqslant 1.0$ 为止,可以得到一系列的状态。

对于 Y 轴方向的一维搜索也可同样进行

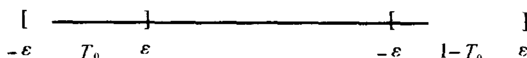


图 2

②二维边界状态算法

i)首先将搜索范围定在 X 轴。以①中的方法进行一维搜索,对每一个状态,转 ii)。

ii)对 i)中某一给定状态,以①中的方法对 Y 轴进行一维搜索,对每一状态,加上 i)中的特定状态,惟一确定了一种状态组合,每种状态组合确定了四个小正方形中心的坐标。由于四个小正方形的位置均已确定,可对所有的点的坐标进行推断,看其是否落在四个小正方形的区域内,记录该种状态组合下所能覆盖的点的个数及各点的下标。

iii)转 i),继续进行搜索,直到满足搜索结束的条件为止。

iv)对所有的状态组合,找出其能覆盖的点数的最大值,则该种状态组合即确定了最优解及这时的最优点组合。



从上面的算法描述可知,程序的内、外循环次数都是 n,因此,搜索次数为 n^2 次。根据钻井的实际情况,当 n 比较小的时候,此算法比“穷举法”省时。

③一维搜索算法的证明

如图 2。初始状态下左区间 L 的右端与 X 轴原点重合。设在任一状态 S_i 下,包括在区间对$(-\varepsilon+x(T_0),\varepsilon+x(T_0))$和$(1-\varepsilon+x(T_0),1+\varepsilon+x(T_0))$内的点数为 M,设在左区间$(x(T_0)-\varepsilon,x(T_0)+\varepsilon)$之处且距离该区间右端点最近的一个点为 Q,Q 与左区间右端点的距离为 Δx。显然,当线段 L 向右移动时,右区间$(1-\varepsilon+x(T_0),1+\varepsilon+x(T_0))$内的点数不可能增加(因所有的点均已归一到$[0,1]$区间内),要使 M 增加,必然要使左区间内的点数增加,因此整个区间对至少要右移 Δx,由 S_i 状态变为 S_{i+1} 状态,这样点数 M 才有可能增加。

所以此种算法不会遗漏最大点数。

④二维搜索算法的证明

反证:假设此算法会遗漏最优值,设最优状态点为 S。那么,此最优状态点 S 必然位于 X 轴上相邻的某两个边界状态点 X_i 与 X_{i+1} 之间。若 S 位于 X_i 或 X_{i+1} 之上(不妨设位于 X_i 上),那么,由于前面我们已经证明当 X_i 固定,仅沿 Y 轴正向一维搜索时,我们可以获得最优值,故此时最优状态点 S 可以被搜索到。从而,最优状态点 S 必然位于 X_i 与 X_{i+1} 内部(不含 X_i 与 X_{i+1})。

当首先沿 X 轴正向搜索时,那么,从 X_i 到状态 S,由于 Y 轴没有移动,所以,同上述一维搜索原理,方框内的点只有可能移出,而不可能有点移入。故 $M(X_i)\geqslant M(S)$,则 S 不是最优状态点,矛盾!

所以假设不成立,即此算法不可能遗漏最优值。

⑤问题一的计算结果

对于题目给出的数据,采用“穷举法”和“二维边界状态法”进行计算,在 Pentium-120/16M 的计算机上运行结果如下:

表 2

	穷举法	二维边界状态法
点数	4	4
点	2,4,5,10	2,4,5,10
偏移量	(0.37,0.47)	(0.36,0.46)
时间	1.04 秒	0.02 秒

可见,本题的一个最优方案为:沿 X 轴正向平移 0.37(或 0.36)单位,再沿 Y 轴正向平移 0.47(或 0.46)单位。

可见,在点不太多的情况下,“二维边界状态法”明显优于“穷举法”。

2. 问题二的建模和求解

i)取步长为 1 度,将连续变化的角度“离散化”,根据对称性,只需在 $\theta \in [0°, 90°]$范围内等距搜索。对于每一个 θ,将 n 个点的坐标进行旋转变换:

$$a_i' = a_i\cos\theta + b_i\sin\theta$$

$$b_i' = -a_i\sin\theta + b_i\cos\theta$$

ii)对 n 个点的坐标进行“归一化”。由于 $\theta \in [0°, 90]$,故 $a_i \geqslant 0$;如果 $b_i < 0$,则 $b_i = |b_i|$,即将 X 轴下方的点作对称变换。

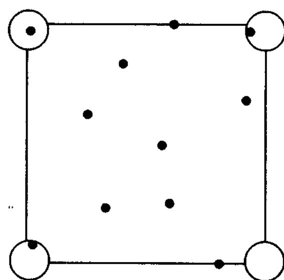


图 3

iii)在欧氏距离意义下,进行二维的格点穷举搜索,找出最优解。

这样,我们得到了一个最优解。

表 3

步长	搜索范围	点数	点序号	旋转角度	偏移量
1°	[0°,90°]	6	1,6,7,8,9,11	45°	(0.77,0.12)

根据最优解的大致范围,我们可以逐步缩小搜索范围,减小步长,以求得更精确的解。

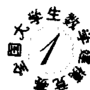


表 4

步长	搜索范围	点数	点序号	旋转角度	偏移量
0.1°	[44°,46°]	6	1,6,7,8,9,11	44.4°	(0.77,0.12)
0.01°	[44.3°,44.5°]	6	1,6,7,8,9,11	44.40°	(0.77,0.12)

可见,本题的一个最优方案为:坐标系绕原点逆时针旋转 44.40°,沿新坐标系下 X 轴正向平移 0.77 单位,再沿 Y 轴正向平移 0.12 单位。

对于这个问题,原网格大小为 10×10 单位,根据估算,只要旋转角步长 $\Delta\theta<0.014°$,则旋转造成的坐标偏移量 $\Delta L<0.05$(精度的一半),不影响结果的精确性。故步长为 0.01°时,求得的解是足够精确的。

3. 问题三的分析和论证

(1)采用第一种距离定义

在不考虑旋转的情况下,可以对所有的点进行归一化处理,从而可导出下列定理:若网格的横向和纵向距离固定时所有旧井可利用的充要条件为:

定理一 若网格的横向和纵向距离固定,则当且仅当任意两点 P_i,P_j,均满足下列条件:

①$|x(P_i)-x(P_j)|\in(0,2\varepsilon)\cup(1-2\varepsilon,1+2\varepsilon)$

②$|y(P_i)-y(P_j)|\in(0,2\varepsilon)\cup(1-2\varepsilon,1+2\varepsilon)$

时所有的旧井均可利用($x(P_i)$是 P_i 的横坐标,$y(P_i)$是 P_i 的纵坐标)。

证明

必要性:

若所有的矿井均可利用,即可以用四个边长为 2ε,中心分别位于一个边长为 1 的大正方形的 4 个顶点的小正方形来覆盖$(0,1)\times(0,1)$区域内的所有点,如图 4。显然任两点之间的关系必然为下列三种情形之一:

①两点位于同一小正方形内。显然有$|x(P_i)-x(P_j)|\in(0,2\varepsilon)$且$|y(P_i)-y(P_j)|\in(0,2\varepsilon)$;

②两点位于相邻的两个小正方形内。若两个小正方形的中心位于同一水平线处,则有$|x(P_i)-x(P_j)|\in(1-2\varepsilon,1+2\varepsilon)$且$|y(P_i)-y(P_j)|\in(0,2\varepsilon)$;若两个小正方形的中心位于同一竖直线上,则有$|x(P_i)-x$

$(P_j)| \in (0,2\varepsilon)$且$|y(P_i)-y(P_j)| \in (1-2\varepsilon,1+2\varepsilon)$；

③两点位于处于对角线位置的两个小正方形内，则显然有$|x(P_i)-x(P_j)| \in (1-2\varepsilon,1+2\varepsilon)$且$|y(P_i)-y(P_j)| \in (1-2\varepsilon,1+2\varepsilon)$。

因此任意两点的坐标必然满足条件一。

充分性：

①在满足条件一的情况下，对于一维的情况，以 X 轴为例，将所有的 n 个点在 X 轴上投影，考虑到由于某些点的 X 坐标相同而导致的重合情况，可得 m 个投影点，且 $m \leqslant n$。

设这 m 个点从左到右依次为：$x_1, x_2, \cdots, x_m$。同时根据条件一，对于 x_1 而言，其余的 $m-1$ 个点可分为 A，B 两类：(1)对任意 $x_i \in A$，$x_i - x_1 \leqslant 2\varepsilon$；(2)对任意 $x_j \in B$，$1-2\varepsilon \leqslant x_j - x_i \leqslant 1+2$。

设 A 中最大的元素为 x_k，B 中最小的元素为 x_{k+1}。对于 B 中的任意两点 x_{i1}, x_{i2}，根据已知条件应有$|x_{i1}-x_{i2}| \in [0,2\varepsilon] \cup [1-2\varepsilon,1+2\varepsilon]$，同时由于 $x_{i1}, x_{i2} \in B$，∴又有$|x_{i1}-x_{i2}| \leqslant [0,4\varepsilon]$。

根据假设条件 $\varepsilon \ll 1$，∴ $4\varepsilon \ll 1-2\varepsilon$。又∵ $4\varepsilon > 2\varepsilon$，∴ $|x_{i1}-x_{i2}| \in [0,2\varepsilon]$。

由 x_{i1}, x_{i2}的一般性，可用两个宽度均为 2ϵ 的闭区间 $C_1[a_1,a_2]$和 $C_2[b_1,b_2]$分别覆盖 A，B 两个集合。

同理还可证明 A 中任一点 x_i 与 B 中任一点 x_j 均满足$|x_i - x_j| \in [1-2\varepsilon,1+2\varepsilon]$。

设区间 C_1，C_2 的中心分别为 O_1，O_2。令 $a_1 = x_1$，$b_2 = x_m$，根据 A，B 中元素的特征可知它们仍然可被$[a_1,a_2]$和$[b_1,b_2]$分别覆盖。

∵ $b_2 - a_1 \in [1-2\varepsilon,1+2\varepsilon]$，而 $a_1 = x(O_1)-\varepsilon$，$b_2 = x(O_2)+\varepsilon$，

∴ $x(O_2)-x(O_1)+2\varepsilon \in (1-2\varepsilon,1+2\varepsilon)$

∴ $x(O_2)-x(O_1) \in [1-4\varepsilon,1]$。

再令 $a_2 = x_k$，$b_1 = x_{k+1}$，同样可知这种情况也是可行的。

∵ $b_1 - a_2 \in [1-2\varepsilon,1+2\varepsilon]$，

而 $a_2 = x(O_1)+\varepsilon$，$b_1 = x(O_2)-\varepsilon$；

∴ $x(O_2)-x(O_1)-2\varepsilon \in (1-2\varepsilon,1+2\varepsilon)$；

∴ $x(O_2)-x(O_1) \in [1,1+4\varepsilon]$。



由于 O_1,O_2 是连续变化的,而根据上述情况可知 $[1-4\varepsilon,1]$,$[1,1+4\varepsilon]$ 均在 $x(O_2)-x(O_1)$ 的可行解范围内。根据连续函数的介值定理,必然可以找到一对区间 C_1,C_2,覆盖了 X 轴上所有的投影点,且 C_1,C_2 的中心距离($x(O_2)-x(O_1)$)为 1。

②同样道理可以在 Y 轴上找出中心相距为 1,宽度为 2ε 的一对区间 C_3、C_4,使之覆盖了 Y 轴上的所有的投影点。

③结合 i),ii),可取 C_1,C_2 及 C_3,C_4 这四个二维平面上的条形区域的交,得四个边长为 2ε 的小正方形区域 $C_1\cup C_3,C_1\cup C_4,C_2\cup C_3,C_2\cup C_4$,它们覆盖了所有的给定点,证毕。

然后,我们根据定理一,从点的分布结构出发,引入"角度离散化"搜索法,将该问题转化为平移问题解决,可采用以下"组合算法"进行判断:

组合算法:

①对于所有的点,进行二元组合,共有 C_n^2 种组合方式。

②对每种组合的点对 (P_i,P_j),同样采用"角度离散化"旋转搜索法,寻找旋转角 θ_{ij} 的范围,使得 P_i,P_j 满足定理一。一旦 θ_{ij} 的范围为空集,则说明 n 口旧井不可能全部利用。转到④。

③若任意 θ_{ij} 的范围都不为空集,则判断 C_n^2 个集合 θ_{ij} 的交集是否为空集。如果为空,则说明 n 口旧井也不可能全部利用。转到④;否则,说明 n 口旧井可以全部利用

④退出程序。

定义一:

θ_{ij}:指对点组合 (P_i,P_j) 采用"角度离散化"旋转搜索法获得的、使得 P_i,P_j 满足定理一的旋转角的范围。

定理二:

在 n 个点 $P_i(i=1,2,\cdots,n)$ 中,如果

$$\bigcap_{i=1}^{n}\bigcap_{j=1}^{n}\theta_{ij}\neq\varnothing$$

则这 n 个点能完全覆盖,即 n 口旧井均可以被利用。

证明

必要性:

若 n 口旧井均可以被利用,即 n 个点可以被全部覆盖。假设此时条件二不被满足,由于 θ 的步长精度足够小,可以认为对任意的 $\theta\in(0,2\pi)$,均无法使任意两点的坐标均满足条件一,从而由条件一的充分必要性可导出无论怎样旋转和平移坐标,均无法使所有的点均被覆盖(∵总有至少两点无法满足条件一)。这与已知假设"n 个点可以被全部覆盖"相矛盾,故假设不成立,必要性得证。

充分性:

令 $\bigcap_{i=1}^{n}\bigcap_{j=1}^{n}\theta_{ij}=A.$

∵ $\bigcap_{i=1}^{n}\bigcap_{j=1}^{n}\theta_{ij}\neq\varnothing$

∴ 至少存在某一 $\theta\in A$,只要将初始的直角坐标系旋转 θ 后,就可以使任意两点都满足条件一,从而说明所有的点均可被覆盖。证毕。

六、模型的评价与改进方向

1.模型的评价

本文讨论了钻井布局模型,采用了"归一化"的简化手段,给出了钻井布局的最优方案。我们给出的"二维边界状态法",在坐标轴不旋转、第一种距离定义下能很快地找到最优解。在坐标轴可以旋转以及采用欧氏距离的条件下,我们采用"角度离散化"和"逐步逼近"的方法,也能找到符合精度要求的最优解。对于问题三,我们选取了第一种距离定义方式,给出了判定 n 口旧井均可利用的定理二。

2.模型的改进方向

在第一种距离定义方式下,我们找到了定理一这样一个判据,因此设计出了"二维边界状态法",减小了时间的耗费。但是,在欧氏距离定义方式下,由于时间关系,我们找不到类似的判据,因此不能设计出更为有效的算法。但我们设想可以通过某些方法对搜索算法进行局部优化。例如可以把解决第一问的"二维边界状态法"推广到第二问当中,用逐点搜索代替网格搜索。如果有充足的时间,我们将考虑是否可能用极坐标来解决有关欧氏距离的问题,因为极坐标系中的(r,θ)与直角坐标系中的$(x,$

y)必然存在某种对应关系,从而使我们可能用(r,θ)平面中的二维逐点搜索来替代(x,y)平面中的二维逐点搜索。还有可能通过某些降低精度的近似算法来优化算法的时间复杂度。

点 评

本文重点研究了棋盘距离下的钻井布局问题。作者采用去除整数部分的办法,将井点“归一”到一个单位正方形中,应用多种搜索技巧设计了能找到最大可利用点数的算法,给出了算法有效性分析,并证明了相应的定理。文章的不足之处之一是未明确给出问题的数学模型。钻井布局要研究的问题虽然本身较为清楚,但仍应写出相应的数学模型。此外,距离可有多种不同的定义方式,对欧氏距离等常用距离下的情况分析,作者也研究的较少。

本文获 1999 年全国大学生数学建模竞赛全国一等奖。

DNA 序列的分类

余　杭　韩轶平　刘　威

摘　要

本文对 A 题中给出的 DNA 序列分类问题进行了讨论。从“不同序列中碱基含量不同”入手建立了欧氏距离判别模型，马氏距离判别模型以及 Fisher 准则判定模型；又从“不同序列中碱基位置不同”入手建立了利用序列相关知识的相关度分类判别算法，并进一步研究了带反馈的相关度分类判别算法。对于题中所给的待分类的人工序列和自然序列，本文都一一作了分类。接着，本文对其他各种常见的分类算法进行了讨论，并着重从分类算法的稳定性上对几种方作了比较。

一、问题的重述

1. 背景

人类基因组中的 DNA 全序列是由 4 个碱基 A，T，G，C 按一定的顺序排成的长约 30 亿的序列。除了这 4 种碱基以外，人们对它包含的内容知之甚少。尽管如此，人类还是发现了 DNA 序列中的一些规律和结构。例如，在全序列中有一些适用于编码蛋白质的序列片断，即由这 4 个碱基组成的 64 种不同的三核苷酸，其中大多数用于编码构成蛋白质的 20 种氨基酸。又例如，在不用于编码蛋白质的序列片断中，A 和 T 的含量特别多些，于是以某些碱基特别丰富作为特征去研究 DNA 序列的结构也取得了一些结果。此外，利用统计方法还发现序列的某些片断之间具有相关性，等等。这些发现让人们相信 DNA 序列中存在着局部和全局性



的结构，充分发掘序列的结构对理解 DNA 全序列是十分有意义的。目前在这项研究中最普通的思想是省略序列的某些细节，突出特征，然后将其表示成适当的数学对象。这种被称为粗粒化和模型化的方法往往有助于研究规律性的结构。

2. 问题

有 20 个已知类的人工制造的样本序列，其中序列标号 $A_1 \sim A_{10}$ 为 A 类，$A_{11} \sim A_{20}$ 为 B 类。要求从中提取特征，构造分类方法，并用这些已知类别的序列，衡量分类方法是否足够好。然后对另外未标明类别的 20 个人工 DNA 序列（$A_{21} \sim A_{40}$）和 182 个自然 DNA 序列（标号 $N_1 \sim N_{182}$）进行分类。

二、模型的条件和假设

1. 题目数据中的错误

在题目所给定的数据中，我们发现了有几组数据存在着不同程度的错误，在数据文件中出现了 s, n, w, r, y 五种不代表碱基的字母。对于这样的情况，人们认为存在两种处理方法：第一，这些含有无意义字母的碱基序列作为不可判类处理；第二，考虑到在一长度很大的碱基序列中出现很少量的错误不会影响到序列的全局特征，可以采用将无意义字母简单删去的方法来处理。本文采用了后一种方式。

2. 正态分布假设

对于特定的一种生物我们认为其 DNA 序列各种碱基的含量服从正态分布。

三、符号约定

n_a：任一给定序列中碱基 A 的百分含量；

n_g：任一给定序列中碱基 G 的百分含量；

n_t：任一给定序列中碱基 T 的百分含量；

n_c：任一给定序列中碱基 C 的百分含量；

G_i：由某些具有相同属性的个体组成的类。

四、问题的分析和解答

1.概述

根据题意，我们首先要提取出一个序列的特征，然后给出它的数学表示，最后选择并构造基于这种数学表示的分类方法。对于任意一个DNA，我们认为，反映该序列特征的方面有两个：

①碱基的含量，反映了该序列的内容；

②碱基的排列情况，反映了该序列的形式。

2.基于碱基含量特征分类的模型

首行，我们考虑采用序列中的A,G,T,C的含量百分比作为该序列的特征。这样抽取特征的方法具有其生物学的意义。前面提到过，在不用于编码蛋白质的序列片断中，A和T的含量特别多些，因此以某些碱基特别丰富作为特征去研究DNA序列的结构是具有可行性的。将序列中的A,G,T,C的含量百分比分别记为n_a,n_g,n_t,n_c，则得到一组表征该序列特征的四维向量(n_a,n_g,n_t,n_c)。考虑到n_a,n_t,n_g,n_c线性相关$(n_a+n_g+n_t+n_c=1)$，所以我们采用简化的三维向量(n_a,n_t,n_g)来进行计算，对于标号为i的序列，记它的特征向量为X_i。显然，任意序列的特征向量与一个三维空间的点对映。

一个一般的判别问题为：设有k个类别$G_1,G_2,\cdots,G_k$，对任意一个属于G_i类样品x，其特征向量X的值都可以获得。现给定一个由已知所属类别的一些样品$x_1,x_2,\cdots,x_n$组成的学习样本，要求对一个来自这k个类别的某样品x，根据其特征向量X的值作出其所属类别的判断。

在本题DNA序列分类中，$k=2,G_1=A,G_2=B$，特征向量X是三维的。学习样本共包含$n=20$个样本，其中10个属于A，10个属于B。我们分别采用了欧氏距离(Euclid)分类模型，马氏距离(Mahalanobis)分类模型和Fisher判别模型来对序列样本分类。

(1)欧氏距离(Euclid)分类模型

在欧氏距离(Euclid)分类模型中，把每个样本视为三维空间的一个



点,以其到不同集合几何中心的欧氏距离作为判据。

具体的算法如下:

①计算属于 A 类的 10 个 点的集合与属于 B 类的 10 个样本点的集合各自的几何中心:

$$C_A = \frac{1}{10}\sum_{i=1}^{10} X_i \qquad C_B = \frac{1}{10}\sum_{i=11}^{20} X_i$$

②对于给定的样本点 X_i,分别计算该点到 C_A 的欧氏距离 $D_A = |X_i - C_A|$,以及该点到 C_B 的欧氏距离 $D_B = |X_i - C_B|$。

③判别准则如下:

a.若 $D_A < D_B$,则将 X_i 点判为 A 类;

b.若 $D_A > D_B$,则将 X_i 点判为 B 类;

c.若 $D_A = D_B$,则将 X_i 点判为不可判类。

用上述算法对已知学习样本 $A_1 \sim A_{20}$进行分类,结果是除了 A_4 被错误地分到 B 类外,其余的 19 样本全部正确,分类准确率达到 95%。

用上述算法对未知的人工序列 $A_{21} \sim A_{40}$进行分类,得到的结果是:

A 类:22,23,25,27,29,30,32,34,35,36,37,39

B 类:21,24,26,28,31,33,38,40

用上述算法对未知的自然序列 $N_1 \sim N_{182}$进行分类,得到的结果见附录。

用欧氏距离作为判据虽然简便直观,但存在着明显的缺陷:从概率统计的角度来看,用欧氏距离描述随机点之间的距离并不好。因此当待分类样本是随机样本,具有一定的统计性质时,这个模型并不能很好地描述两个随机点之间的接近程度。

(2)马氏距离(Mahalanobis)分类模型

为了克服采用欧氏距离时的缺陷,采用了马氏距离来代替欧氏距离。

改进后的算法如下:

设:三维总体 G 的均值为 $\mu = (\mu_1, \mu_2, \mu_3)^T$,协方差矩阵为非奇异阵 $V_{3\times 3}$,则三维样本 X 到总体 G 的马氏距离为:

$$d_m(X,G) = \sqrt{(X-\mu)^T V^{-1}(X-\mu)},$$

其中,未知的 μ 可用学习样本的均值来代替,协方差矩阵 V 可用学习样

本的样本协方差矩阵来代替。

将马氏距离用于判别模型，遵循判据如下：

a.若 $d_m(X,A) < d_m(X,B)$，则判定 X 为 A 类；

b.若 $d_m(X,A) > d_m(X,B)$，则判定 X 为 B 类；

c.若 $d_m(X,A) = d_m(X,B)$，则判定 X 为不可判类。

用上述算法对已知学习样本 $A_1 \sim A_{20}$ 进行分类，结果是除了 A_4 被错误的分到 B 类外，其余的 19 个样本全部正确，分类准确率达到 95%。

用上述算法对未知序列 $A_{21} \sim A_{40}$ 进行分类，得到的结果是：

A 类：22,23,25,27,29,30,32,33,34,35,36,37

B 类：21,24,26,28,31,38,39,40

用上述算法对未知的自然序列 $N_1 \sim N_{182}$ 进行分类，得到的结果见附录。

(3)Fisher 准则分类模型

值得指出的是，在多维空间里分类的方法不仅仅是距离分类法一种，常用的 Fisher 分类法就是另一种基于几何特性的分类法。在距离判别模型中，三维空间的样本 X 被映射为一维的距离 d，依其来作判断。Fisher 分类法的思想也是把三维空间的样本映射为一维的特征值 y，并依据 y 来进行判断。具体的作法是先引入一个与样本同维的待定向量 u，再将 y 取 X 坐标的线性组合 $y = u^T x$。而 u 的选取要使同一类别产生的结果尽量聚拢，不同类别产生的 y 尽量拉开。这样，我们便可将样本 X 到某一类 G 的距离定义为 $y = u^T x$ 与 $yc = u^T c$ 之间的欧氏距离：

$$L(X,G) = |y - yc| = |u^T(x - c)|$$

其中，c 为 G 的几何中心。

Fisher 分类的判据为：

a.若 $L(X,A) < L(X,B)$，则判定 X 为 A 类；

b.若 $L(X,A) > L(X,B)$，则判定 X 为 B 类；

c.若 $L(X,A) = L(X,B)$，则判定 X 为不可判类。

根据对 u 的要求，Fisher 提出了一个比较有效的选择算法，利用该算法，我们可以从学习样本中获得：

$u = (0.3365, -0.087, 0.9377)^T$



$$L(X,A)=|0.3365(n_a-0.2860)-0.087(n_t-0.1550)+0.9377(n_g-0.3830)|$$

$$L(X,B)=|0.3365(n_a-0.2940)-0.087(n_t-0.5010)+0.9377(n_g-0.1010)|$$

用上述算法对已知样学习样本 $A_1 \sim A_{20}$进行分类,结果仍然是除了 A_4 被错误的分到 B 类外,其余的 19 个样本全部正确,分类准确率达到 95%。

对于未知序列 $A_{21} \sim A_{40}$进行分类,得到的结果是:

A 类:22,23,25,27,29,34,35,36,37

B 类:21,24,26,28,30,31,32,33,38,39,40

用上述算法对未知的自然序列 $N_1 \sim N_{182}$进行分类,得到的结果见附录 C。

(4)三种距离分类模型的比较

这三种模型在分类结果上有一定的区别,对于序列 A_{30},A_{32},A_{33}及 A_{39},三种方法给出了不同结果,见下表:

	欧氏距离法	马氏距离法	Fisher 准则法
30	A	A	B
32	A	A	B
33	B	A	B
39	A	B	B

对于这种情况,我们提出一个联合判定准则:对于任一个序列,当三种分类法结果完全一致时,认为它判别有效;若不然,当三种分类结果不一致时,认为该序列为不可判类。

另外对于三种方法都无法正确分类的 A_4 序列,可认为这是一个异常情况,不影响算法的性能。

3.基于碱基位置特征分类的模型

虽然上述采用碱基 A,T,G,C 在 DNA 序列里的含量作为该序列特征的方法有一定的生物学意义,并且在 DNA 序列的分类中获得了比较理想的结果。但是,用这种方法抽取特征,没有充分体现碱基排列的信息

量,仅仅考虑碱基含量并没有体现碱基序列中的排列情况。例如,序列(ATGC)与序列(CGTA)有着相同的碱基含量,他们的特征向量是完全一样的,并不能体现他们在排列结构上的不同。因此,直接从序列本身的碱基排列顺序来考察序列就成为一种更加合适的提取特征的方式。因此,我们采纳了数值序列中的相关性分析思想设计了算法。

通常任意两个数值序列的相关性都是通过这两个序列的相关函数来刻画的。由于本题中的 DNA 序列是非数值的序列,同时无法将碱基按通常的方式进行数值化,因而刻画任意两个序列的相关程度的变量需要重新定义。

(1)定义一:相关运算"$\otimes$"。

对于任意碱基 m 和 n,相关运算"$m\otimes n$"的值由下表定义:

$\otimes$	A	G	T	C
A	1	0	0	0
G	0	1	0	0
T	0	0	1	0
C	0	0	0	1

(2)定义二:哑元 O

除四个碱基外,我们另行定义一个哑元 O,规定任意碱基与哑元作相关运算的结果都为 0。

(3)定义三:序列的延拓

对于任意一个长度为 N 的序列 $A=\{A_0,A_1,\cdots,A_n\}$,定义它的延拓为如下一个无限序列:

A_j^+:当 $0\leqslant j<N$ 时,$A_j^+=A_j$;

当 $-\infty<j<0$ 及 $N\leqslant j<\infty$ 时,$A_j^+=O$。

即在该序列的左右两端均用哑元 O 填充。

(4)定义四:序列的相关度

对于任意的两个序列 A_N,B_M,定义序列 A 和序列 B 的相关序列 S_i 为:

$$S_i = \sum_{k=0}^{\infty} A_{k+2-i}^{+} \otimes B_k^{+} \qquad (0 \leqslant i \leqslant n+m-1)$$

定义序列 B 对序列 A 的相关度为：

$$S = \min_{(0\leqslant i\leqslant n+m-1)}$$

例如对于序列 $A=\{T,C,T\}$ 与序列 $B=(A,G,T,C,T,C)$，相关序列及相关度的计算步骤如下：

第一项：$S_0 = A_2 \otimes B_0 = T \otimes A = 0$

…	A_{-1}^+	A_0^+	A_1^+	A_2^+	A_3^+	A_4^+	A_5^+	A_6^+	A_7^+	A_8^+	…
…	O	T	C	T	O	O	O	O	O	O	…
…	O	O	O	A	G	T	C	T	C	O	…
…	B_{-3}^+	B_{-2}^+	B_{-1}^+	B_0^+	B_1^+	B_2^+	B_3^+	B_4^+	B_5^+	B_6^+	…

第二项：$S_1 = A_1 \otimes B_0 + A_2 \otimes B_1 = T \otimes G + C \otimes A = 0$

…	A_{-2}^+	A_{-1}^+	A_0^+	A_1^+	A_2^+	A_3^+	A_4^+	A_5^+	A_6^+	A_7^+	…
…	O	O	T	C	T	O	O	O	O	O	…
…	O	O	O	A	G	T	C	T	C	O	…
…	B_{-3}^+	B_{-2}^+	B_{-1}^+	B_0^+	B_1^+	B_2^+	B_3^+	B_4^+	B_5^+	B_6^+	…

第三项：$S_2 = A_0 \otimes B_0 + A_1 \otimes B_1 = T \otimes T + G \otimes C + A \otimes T = 1$

…	A_{-3}^+	A_{-2}^+	A_{-1}^+	A_0^+	A_1^+	A_2^+	A_3^+	A_4^+	A_5^+	A_6^+	…
…	O	O	O	T	C	T	O	O	O	O	…
…	O	O	O	A	G	T	C	T	C	O	…
…	B_{-3}^+	B_{-2}^+	B_{-1}^+	B_0^+	B_1^+	B_2^+	B_3^+	B_4^+	B_5^+	B_6^+	…

以下类推得(表略)：

第四项：$S_3 = A_0 \otimes B_1 + A_1 \otimes B_2 + A_2 \otimes B_3 = T \otimes C + C \otimes T + T \otimes G = 0$

第五项：$S_4 = A_0 \otimes B_2 + A_1 \otimes B_3 + A_2 \otimes B_4 = T \otimes T + C \otimes C + T \otimes T = 0$

第六项：$S_5 = A_0 \otimes B_3 + A_1 \otimes B_4 + A_2 \otimes B_5 = T \otimes C + C \otimes T + T \otimes C = 0$

第七项：$S_6 = A_0 \otimes B_4 + A_1 \otimes B_5 = C \otimes C + T \otimes T = 2$

第八项：$S_7 = A_0 \otimes B_5 = T \otimes C = 0$

第八项：$S_7 = A_0 \otimes B_5 = T \otimes C = 0$										
…	A_{-8}^+	A_{-7}^+	A_{-6}^+	A_{-5}^+	A_{-4}^+	A_{-3}^+	A_{-2}^+	A_{-1}^+	A_0^+	A_1^+	…
…	O	O	O	O	O	O	O	O	T	C	…
…	O	O	O	A	G	T	C	T	C	O	…
…	B_{-3}^+	B_{-2}^+	B_{-1}^+	B_0^+	B_1^+	B_2^+	B_3^+	B_4^+	B_5^+	B_6^+	…

两序列的相关度为 $S = \max\{S_i\} = S_5 = 3$；

(5)**定理一**：任意给定三个序列 S, A, B，若 A 与 S 的相关度大于 B 与 S 的相关度且 B 与 A 等长，则 A 与 S 属同一类的可能性大于 B 与 S 属同一类的可能性

(6)基于相关度的分类算法

利用上述概念，我们构造了一个基于相关度的分类算法如下：

①对于序列 $A_{21} \sim A_{40}$，$N_1 \sim N_{182}$ 中的任意一个序列，将其与序列 $A_1 \sim A_{20}$ 中的每一个依次作求相关度的运算，结果记为 $SS1, SS2, SS3, \cdots, SS20$；

②对于前 10 个相关度，求出它们的平均相关度 $SA = (SS1 + SS2 + \cdots + SS10)/10$，并定义其为与 A 类的相关度；

③对于后 10 个相关度，求出它们的平均相关度 $SB = (SS11 + SS12 + \cdots + SS20)/10$，并定义其与 B 类列的相关度；

④记 $W = SA/SB$，根据定理一，判别依据为：

若 $W > 1$，则将 X 点判为 A 类；

若 $W < 1$，则将 X 点判为 B 类；

若 $W = 1$，则将 X 点判为不可判类。

⑤W 可作为衡量该序列分类的可信性的一个标准。显然当 W 越接近于 1，该序列与 A 类的相关性和与 B 类的相关性区别就越小，分类结果就越不可信；反之，W 与 1 差的越远，该序列与 A 类的相关性和 B 类的相关性区别就越小，分类结果就越可信。这个变量对我们下面带有反馈的相关度分类算法具有重要的意义。

（需要说明的是，A 类的 10 个样本中的 9 个长度为 111，一个为 110，而 B 类的 10 个样本长度均为 110，近似满足定理的一条件，故可按照该



定理设计出此算法。)

用上述算法对已知样本 $A_1 \sim A_{20}$ 进行分类,得到的结果是分类完全正确,A,B 类可以完全分开,准确率达到 100%。

对于未知序列 $A_{21} \sim A_{40}$ 进行分类,得到的结果是:

A 类:22 23 25 27 29 34 35 36 37

B 类:21 24 26 28 30 31 32 33 38 39 40

用上述算法对未知的自然序列 $N_1 \sim N_{182}$ 进行分类,得到的结果见附录。

(7)相关度分类算法的改进——带有反馈的分类算法

上述的相关度分类算法是一次性学习过程,学习的过程只体现在学习样本的过程中,而在对未知样本分类的过程中没有对已分类情况作出修正,即是属于无反馈型的学习。然而,采用反馈型的学习过程会有更好的分类结果。一般说来,带反馈的算法以神经网络算法最具有代表性。但对于一般的分类算法而言,可以采用多次反复分类的办法来实现反馈的目的。针对上述的相关度分类算法,我们设计了如下带反馈的相关度分类算法:

①对全部 182 个样本进行相关度分类;

②计算全部 182 个 W 的值;

③在所有被判为 A 类的待分类序列中,取出 W 值最大的一个,作为标准学习样本,加入到 A 类的标准样本中(若有多个,则全部加入到 A 类中,若无被判为 A 类的序列,则保持 A 类标准学习样本不变);

④在所有被判为 B 类的待分类序列中,取出 W 值最小的一个,作为标准学习样本,加入到 B 类的标准样本中(若有多个,则全部加入到 B 类中,若无被判为 B 类的序列,则保持 B 类标准学习样本不变);

⑤重复对剩余的待分类序列进行相关分类,并按上述步骤不断扩充标准学习样本,直至全部的待分类序列都被加入到标准学习样本中。

我们用新算法编程对 182 个序列进行了重新分类,得到了不同于原无反馈分类算法的结果,而且新的分类结果的 W 值明显与 1 离开得更大,这使我们有理由相信,反馈对算法的性能有一定的改进。



五、进一步研究的问题

1.基于生物学的特征抽取

上述两种特征抽取方法更多地是从纯数学眼光来研究序列的特征。除此之外，我们还可以考虑 DNA 序列在生物学意义下的数学特征。

一个比较容易考虑到的方面便是三联体在 DNA 序列中的出现。由于具有三联体形式的遗传密码子对蛋白质的合成具有决定性作用，有理由认为它在序列中的出现体现了该序列的本质特征。题中没有明确地指明所给的序列是全序列还是序列片断，我们无法对三联体在序列中的出现位置进行定位，一种代替的方法是将序列假定为全序列，从第一个碱基开始三个三个一组地划分为密码子，然后统计 64 个密码子的出现概率，形成 64 维的向量。再使用距离分类等模型，或利用生物学的知识先将 64 维向量的某几维合并，降维后再分类。我们编程演算后，觉得该种分类方法比较依赖于密码子的划分，一位碱基的缺失或错位均会造成分类错误，所以必须加以修改，一条思路是尝试将序列移一位或二位再划分密码子，由于时间所限，没有进一步研究。

2.基于人工神经网络的模型

人工神经网络是一种带反馈，自适应的算法，随着计算机的计算速度提高而被广泛应用。对于本题的情况采用神经网络模型是合适的，它可以在给定特征向量的情况下代替一般的距离分类模型。对于基于碱基含量的特征(n_a, n_t, n_g)，我们构造了如下的反向传播算法：

(1)网络简单的分为两层，一层为输入层，有 3 个单元，分别为权重 a, b, c；一为输出层，有 1 个单元，为判别结果；各单元均为 Sigmoid 型函数激励。

(2)设定(a, b, c)的初值为$(0,0,0)$；A 类学习样本的标准输出定为 1；B 类学习样本的标准输出定为 0。

(3)对每一个学习样本，计算 $S = an_a + bn_t + cn_g$ 作输出。

(4)将学习样本的标准输出与 S 相减，所得的差用来指导权重的改变，权重的改变遵从 Widrow-Hoff 准则。



(5)反复学习样本,到权重值稳定收敛。

(6)代入待分类样本,分类。

用上述算法所得到的结果与普通的分类模型没有区别。事实上当权值稳定收敛后,$S = an_a + bn_t + cn_g$ 就是特征空间的一张(超)平面,从这一点来说,人工神经网络模型与一般的距离分类模型得到的结果没有两样。考虑到人工神经网络模型还存在结果对初值有较强敏感性,缺乏选择理想步长的准则和收敛性等问题,在较短的时间内,我们无法较好地解决这些问题,所以我们也没有作进一步讨论。

六、算法的稳定性

前面比较算法的时候,曾多次提到分类算法的稳定性问题。分类算法的稳定性是除了算法的成功率之外的另一较重要的指标。所谓分类算法的稳定性,是指算法在样本发生了轻微变化时作出正确判别的能力。对于本题,是指算法在样本序列发生了轻微的碱基缺失、错位、错排情况时作出正确判别的能力。因为本题要求我们研究的是 DNA 序列粗粒化和模型化的问题,所以分类时是对序列的整体特征进行区分。局部碱基的组成变化应该对算法的分类结果没有影响。我们所提出的几个模型均较好地满足这一点。

七、参考书目

[1]孙乃恩,孙东旭,朱德煦.分子遗传学.南京:南京大学出版社,1996

[2]白其峥.数学建模案例分析.北京:海洋出版社,2000

[3]潘德惠.数学建型的统计方法.沈阳:辽宁科学技术出版社,1986

[4]阎平凡,黄端旭.人工神经网络.合肥:安徽教育出版社,1991

[5]李振刚.分子遗传学概论.合肥:中国科学技术大学出版社,1990

[6]Duane Hanselman, Bruce Littlefield. Mastering MATLAB: a comprehensive tutorial and reference, Prentice Hall, 1996

附录:序列 $N_1 \sim N_{182}$ 的分类结果

欧氏距离分类法

A:1 2 3 4 5 6 7 8 9 11 12 13 14 15 16 17 18 19 20 21 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 44 45 46 47 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 77 78 79 81 82 83 87 88 89 90 91 93 94 95 96 97 98 99 100 101 104 105 106 108 109 111 112 113 115 116 117 118 120 121 123 124 125 126 127 129 132 133 134 135 136 137 139 140 141 142 143 145 146 147 148 149 150 152 153 154 155 157 158 160 164 165 167 168 169 171 172 173 174 175 176 177 178 179 180 181

B:7 10 22 23 43 48 75 76 80 84 85 86 92 102 103 107 110 114 119 122 128 130 131 138 144 151 156 159 161 162 163 166 170 182

马氏距离分类法

A:1 2 3 4 5 6 7 8 9 11 12 13 14 15 16 17 18 19 20 21 22 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 41 42 44 45 46 47 49 50 51 52 53 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 76 78 79 81 82 83 84 85 86 87 88 89 90 91 93 95 96 97 98 99 100 101 103 104 105 106 108 109 111 112 113 115 116 117 118 120 121 123 124 125 126 127 128 129 130 132 133 134 135 136 137 139 140 141 142 143 145 146 147 148 149 150 152 153 154 155 156 157 158 159 160 164 165 167 168 169 171 172 173 174 175 176 177 178 179 180 181

B:10 23 40 43 48 74 75 80 92 94 102 107 110 114 119 122 131 138 144 151 161 162 163 166 170 182

Fisher 分类法

A:3 4 5 8 9 13 14 15 16 17 18 19 20 21 25 27 31 32 33 35 36 38 39 41 42 44 45 47 49 51 52 53 55 58 59 60 61 62 64 66 67 68 69 70 71 73 77 79 82 87 89 90 91 93 98 99 100 101 104 105 106 108 109 111 112 113 115 117 118 120 124 132 134 136 139 141 145 148 150 153 154 155 158 171 172 176

B:1 2 6 7 10 11 12 22 23 24 26 28 29 30 34 37 40 43 46 48 50 54 56 57 63 65 72 74 75 76 78 80 81 83 84 85 86 88 92 94 95 96 97 102 103 107 110 114 116 119 122 123 125 126 127 128 129 130 131 133 135 137 138 140 142 143 144 146 147 149 151 152 156 157 159 160 161 162 163 164 165 166 167 168 169 170 173 174 175 177 178 179



180 181 182

相关度分类法

A:2 7 12 13 14 15 16 19 22 26 32 33 35 36 37 40 46 47 48 49 51 53 54 56 57 59 60 62 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 103 104 106 107 109 111 112 113 114 115 116 117 118 119 136 140 141 142 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 171 173 174 177 179 180 181 182

B:1 3 4 6 8 9 10 11 17 18 20 21 23 24 25 27 28 29 30 31 34 38 39 41 42 43 44 45 50 52 55 57 61 63 64 65 66 67 68 769 70 71 72 73 102 105 108 110 111 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 137 138 139 143 165 169 170 173 176 178

点 评

本文借助计算机进行符号处理的能力来把握 DNA 序列中不同碱基的丰度特征。利用统计学方法来进行 DNA 序列的分类研究。文章首先从碱基含量入手，建立了欧氏距离判别模型、马氏距离判别模型以及 Fisher 准则判别模型，提出了联合判别准则。进而提出通过比较碱基相对位置来进行 DNA 序列分类，提出了相关度分类判别算法，并定为带反馈机制的相关度分类判别算法，在既定的方向上颇具新意地将工作引向深入。本文给出的分类方法不仅对人工数据得到了理想的分类，对 182 个自然数据的分类的正确率也很高。

由于赛题的开放性，给参赛选手很大的创新、发挥空间。不同的选手可以提出各自完全不同的分类方法，如何衡量、评价方法是关键。该文的作者注意到了这一点，文章中给出的评价准则除了分类的准确性外，还有算法的稳定性。文章提出了各种算法，但不是各种算法的简单罗列，而是对这些算法的优缺点有分析、有综合、有比较、有结论。显示出对研究内容和结果的很好地把握。例如，文章提到人工神经网络算法，并给出了为什么放弃此算法的“一家之言”。再如，文章给出的模型和方法主要是基于纯数学的角度进行探讨，但在文章最后作者还不失时机地对编码氨基



酸的三联体进行了初步讨论，对文章的缺陷进行了弥补。这些均说明了他们参赛的 3 天时间里对问题的研究是相当全面和深刻的。

文章内容丰富，条理清楚，层次分明，作为 3 天时间里完成的工作，称得上是一篇佳作。值得一提的是该队在参加 1999 年全国大学生数学建模竞赛中，虽较好地完成了赛题的要求，提出了一个有效算法，并取得了很好的结果，但由于在文章写作方面存在一定的缺陷，无法让读者清楚、全面地了解作者的思想和意图，最终名落孙山。在接下来的时间里他们不气馁，继续在自己的弱项上下功夫，进步很快，连续在 2000 年全大学数学建模竞争中获全国一等奖，在 2000 年及 2001 年美国大学数学建模竞争中均获国际一等奖，取得了优异成绩。

本文获 2000 年全国大学生数学建模竞赛全国一等奖。

钢管购运和管道铺设方案设计

孙 杰 赵 明 葛志兵

摘 要

本文解决的是一个钢管购运和管道铺设方案设计问题,目的是使总费用最小。首先利用动态规划方法求解所有钢厂到各管道节点的最小运费表,并通过分析得出从管道两边节点向中间铺路的方法可以减少铺设费用的结论;以所有钢厂到各管道节点并向不同方向铺设的钢管数量为变量,导出总费用的表达式,把问题化为以总费用为目标函数的非线性规划,用 Matlab 对此进行求解。然后通过钢厂钢管销价和产量上限的微小变化及在新的条件下的求解,分析对总费用和购运计划的影响。最后把模型推广到树形管道,通过变换把它化为多条线形管道,用同样方法求解。由于计算中变量个数的增加使计算复杂度提高,这里提供了一些优化策略。在本文的末尾,讨论了模型的优缺点和实际应用中的改进方向。

本文利用以上算法较好地解决了问题,得到了问题的最优解。对于问题一,解得最小总费用为 127.86 亿元,购运和铺设方案见表 2。对于问题二,分析得出 $S6$ 厂的钢管价格变化对总费用影响最大,$S1$ 厂的产量上限变化对总费用影响最大。对于问题三,解得最小总费用为 140.66 亿元,购运和铺设方案见表 6。

一、问题重述

要铺设一条线形的天然气运输管道,钢管由 7 个钢厂提供,可以由公路、铁路运往铺设地点。钢厂提供的钢管量不能超过它的产量上限,且或

者大于 500 或者为 0。公路、铁路运输费用的计算公式已经给定。在此条件下求出定购运输计划,使总费用最小。并分析钢厂钢管销价和产量上限变化对购运计划和总费用的影响。以及推广模型以解决铺设树形管道这种更一般的情形。

二、模型的假设

(1)公路运输费用为 1 单位钢管每公里 0.1 万元,不足整公里按整公里计算。

(2)购买和运输钢管都是整单位(即为整公里)。

(3)沿管道或者原来有公路或者建有施工公路。

(4)一个钢厂如果承担制造钢管,至少要生产 500 个单位。

(5)钢管可由铁路、公路运往铺设地点。

(6)把"钢厂钢管的销价和产量上限变化对总费用和运购计划的影响"理解为在最优解附近的微小变化对总费用和运购计划的影响。销价最小变化是 1 万元,产量上限的最小变化是 1 个单位。

三、问题分析

铺设一个天然气运输管道(线形或树形),总费用包括购买钢管的费用,运费和铺设时的费用。购买钢管的费用由钢厂的钢管销价和向这个厂订购的钢管数量决定。运费由钢厂向铺设起始点运输的钢管数量和它到此起始点的运输道路决定(由于通过铁路和公路运输,所以并不仅仅由路程决定)。一般情况下铁路运输比公路运输要节省费用(只有在 200 公里以内,公路运输比铁路运输要节省)。对于铺设费用,在假设一的前提下,由一头出发,铺设 x 公里的费用的计算是:$0.1[x+(x-1)+(x-2)+\cdots+2+1]=0.05x(x+1)$,通过比较可以发现:铺设一段管道,从两头往中间铺比从一端向另一端铺要节省费用。再由假设三,铺设时沿管道走的是公路,所以,当管道段较长时,两头铺所节省的费用是比较可观的。比如问题一中,所有管道段都从一头铺的铺设总费用为:$0.05\sum[D_{x,y}\times(D_{x,y}+1)]=12.29$ 亿元。而在最优的铺设方法下(即两头向中间铺的



路程相同),铺设总费用为:$0.05\sum 2\times\left[\frac{D_{x,y}}{2}\times\left(\frac{D_{x,y}}{2}+1\right)\right]=6.16$ 亿元,最优解中的铺设费用应在这两者之间。因为铺设费用的表达式是二次式,所以求解总费用是一个非线性规划问题。

四、符号说明

p_i: 钢厂 S_i 的钢管销售价格;

s_i: 指定期限内钢厂 S_i 能生产钢管的最大数量;

$L_{i,j}$: 从 S_i 运到 A_j,且向左边铺路的钢管数量;

$R_{i,j}$: 从 S_i 运到 A_j,且向右边铺路的钢管数量;

T_i: 从 S_i 运出的钢管总量,要求 $T_i \leqslant s_i$,且 $T_i=0$ 或 $T_i \geqslant 500$;

$F_{i,j}$: 一单位钢管从 S_i 运到 A_j 的最少费用;

$D_{x,y}$: 相邻两点 A_x,A_y 之间的路程;

C_1: 购买钢管的费用;

C_2: 把钢管运送到所有 A_j 的总运费;

C_3: 从 A_j 开始铺设钢管过程中的公路运费;

C: 总费用,$C=C_1+C_2+C_3$;

k_i: S_i 厂钢管价格对总费用的边际影响;

m_i: S_i 厂钢管产量上限对总费用的边际影响。

五、模型的建立与求解

1.问题一的订购和运输计划求解

(1)模型一的建立

由定义得:

从 S_i 运出的钢管总量 $T_i=\sum_{j=1}^{15}(L_{i,j}+R_{i,j})$

购买钢管的费用 $C_1=\sum_{i=1}^{7}(T_i\cdot p_i)$

把钢管运送到所有 A_j 的总运费 $C_2=\sum_{j=1}^{15}\sum_{i=1}^{7}[(L_{i,j}+R_{i,j})\cdot F_{i,j}]$

再设从 A_j 向左铺设的钢管量 $L_j = \sum_{i=1}^{7} L_{i,j}$

向右铺设的钢管量 $R_j = \sum_{i=1}^{7} R_{i,j}$

则铺设钢管过程中的公路运费：

$$C_3 = 0.1 \times \sum_{j=1}^{15} [L_j(L_j + 1)/2 + R_j(R_j + 1)/2]$$

目标函数（总费用）为：$\min(C_1 + C_2 + C_3)$

s.t. $T_i \leqslant s_i$，且 $T_i = 0$ 或 $T_i \geqslant 500$ $(i = 1,2,\cdots,7)$

对相邻两点 A_x, A_y 有 $R_x + L_y = D_{x,y}(x, y = 1,2,\cdots,15)$

问题即转化为对以 $L_{i,j}, R_{i,j}$ 为变量的非线性规划的求解。

(2)先求出从 S_i 运送单位钢管到 A_j 的最少费用 $F_{i,j}$，可用一个类似求最短轨道的动态规划算法，可以仿照 Dijkstra 算法[1]，特殊的是边的权以路程表示，而阶段指标是运费，求解结果见表 1：

表 1　最小运费表（问题一）　　　　单位（万元）

	S_1	S_2	S_3	S_4	S_5	S_6	S_7
A_1	170.7	215.7	230.7	260.7	255.7	265.7	275.7
A_2	160.3	205.3	220.3	250.3	245.3	255.3	265.3
A_3	140.2	190.2	200.2	235.2	225.2	235.2	245.2
A_4	98.6	171.6	181.6	216.6	206.6	216.6	226.6
A_5	38.0	111.0	121.0	156.0	146.0	156.0	166.0
A_6	20.5	95.5	105.5	140.5	130.5	140.5	150.5
A_7	3.1	86.0	96.0	131.0	121.0	131.0	141.0
A_8	21.2	71.2	86.2	116.2	111.2	121.2	131.2
A_9	64.2	114.2	48.2	84.2	79.2	84.2	99.2
A_{10}	92.0	142.0	82.0	62.0	57.0	62.0	77.0
A_{11}	96.0	146.0	86.0	51.0	33.0	51.0	66.0
A_{12}	106.0	156.0	96.0	61.0	51.0	45.0	56.0
A_{13}	121.2	171.2	111.2	76.2	71.2	26.2	38.2
A_{14}	128.0	178.0	118.0	83.0	73.0	11.0	26.0
A_{15}	142.0	192.0	132.0	97.0	87.0	28.0	2.0

(3)用 Matlab 编程求解目标函数，得到最优解为 127.54 亿元，此时从 $S_1, S_2, \cdots, S_7$ 运出的钢管数量为：$T_1 = 800$，$T_2 = 800$，$T_3 = 1000$，



$T_4=0$，$T_5=1336$，$T_6=990$，$T_7=245$。但是由于 $T_7=245<500$，而题目中要求"一个钢厂如果承担制造钢管，至少需要生产 500 个单位"，所以此解不符合条件。

(4)再分两部分进行第二步求解，一是在增加约束 $T_7=0$ 的情况下，二是在增加约束 $T_7 \geqslant 500$ 的情况下分别计算。对于第一种情况，在实际计算中我们发现，只要把 S_7 销售钢管的价格 p_7 变得很高(比如 1000 万元)，那么在最后的结果中肯定不会出现有 S_7 提供钢管的情况，这样求得的结果与增加约束 $T_7=0$ 相同。由 Matlab 解得总费用为 127.86 亿元。对于第二种情况，求得的结果是 127.97 亿元。并且在两种情况下所有的 T_i 都满足条件，故不需要再继续求解。最优解是第一种情况，即当 $T_7=0$ 时的结果。

补充：如果在计算中，第二步求解得到的解仍然有 T_i 不满足条件，再对此 T_i 分几部分进行下一步求解，依次类推。

最优解对应的钢管购运计划如表 2：

表 2 钢管购运和铺设计划(问题一)

	S_1		S_2		S_3		S_4		S_5		S_6		S_7	
	左	右	左	右	左	右	左	右	左	右	左	右	左	右
A_1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A_2	0	0	104	75	0	0	0	0	0	0	0	0	0	0
A_3	0	0	226	0	0	0	0	0	0	282	0	0	0	0
A_4	37	0	95	0	336	0	0	0	0	0	0	0	0	0
A_5	298	0	0	0	0	0	0	0	308	10	0	0	0	0
A_6	184	15	0	0	0	0	0	0	0	0	0	0	0	0
A_7	190	76	0	0	0	0	0	0	0	0	0	0	0	0
A_8	0	0	125	175	0	0	0	0	0	0	0	0	0	0
A_9	0	0	0	0	505	159	0	0	0	0	0	0	0	0
A_{10}	0	0	0	0	0	0	0	0	321	30	0	0	0	0
A_{11}	0	0	0	0	0	0	0	0	270	145	0	0	0	0
A_{12}	0	0	0	0	0	0	0	0	0	0	75	11	0	0
A_{13}	0	0	0	0	0	0	0	0	0	0	199	134	0	0
A_{14}	0	0	0	0	0	0	0	0	0	0	286	335	0	0
A_{15}	0	0	0	0	0	0	0	0	0	0	165	0	0	0
T_i	800		800		1000		0		1366		1205		0	

说明：

①表中表示从 S_i 订购并运输到 A_j 的钢管数量，以及这些钢管向左和向右铺设的数量。

②虽然表中有从 S_i 运到 A_j 的钢管向左和向右的分配量，实际上只要所有的钢管运到 A_j 后，向左和向右的铺设量与钢管的来源无关。

在上表的方案下，问题一的总费用最小为 127.86 亿元。

2. 钢厂销价和产量上限变化的影响分析

(1)钢厂销售价格变化分析

我们把任一钢厂销售价格在原来的基础上增加和减少 1 万元，并用同样方法求得此条件下的最优解 C' 和 C''，计算钢管价格对总费用的边际影响 $k_i = \left|\frac{\partial C}{\partial p_i}\right| = \left(\left|\frac{C'-C}{1}\right| + \left|\frac{C''-C}{1}\right|\right)/2$ 得到：

表 3　钢管价格对总费用的边际影响

钢厂 S_i	S_1	S_2	S_3	S_4	S_5	S_6	S_7
$\lvert C'-C\rvert$ (万元)	800	800	1000	0	1008	1203	0
$\lvert C''-C\rvert$ (万元)	800	800	1000	0	1368	1563	0
钢管价格边际影响 k_i	800	800	1000	0	1188	1383	0

(2)钢厂产量上限变化分析

同样地，我们把任一钢厂的产量上限增加和减少 1 万元，求得最优解 C' 和 C''，计算钢厂产量上限对总费用的边际影响 $m_i = \left|\frac{\partial C}{\partial s_i}\right| = \left(\left|\frac{C'-C}{1}\right| + \left|\frac{C''-C}{1}\right|\right)/2$，结果见表 4：

表 4　钢厂产量上限对总费用的边际影响

钢厂 S_i	S_1	S_2	S_3	S_4	S_5	S_6	S_7
$\lvert C'-C\rvert$ (万元)	103	35	25	0	0	0	0
$\lvert C''-C\rvert$ (万元)	103	35	25	0	0	0	0
产量上限边际影响 m_i	103	35	25	0	0	0	0



①从表 3,表 4 看出:S_1,S_2,S_3 三个厂的钢管全部被购买(由此可以得出对这三个厂的钢管需求量比较高),所以这三个厂的产量上限变化会对总费用产生影响,而其他四个厂在最优解中产量没有达到上限,所以它们的产量上限在较小变化范围内对总费用没有影响。同时,我们得出 S_1,S_2,S_3 三个厂的钢管价格在较小范围内变化并不会影响钢管订购和运输的安排(它对总费用的影响仅在于对购买钢管费 C_1 的影响)。

②S_5,S_6 的钢管价格变化会影响钢管订购和运输的安排(主要是 S_5,S_6 钢管购运的重新分配)。

所以问题二的解答是:S_6 厂的钢管价格变化对总费用影响最大,S_1 厂的产量上限变化对总费用影响最大。

3. 问题三的订购和运输计划求解

问题三中铺设的管道不是一条线,而是一个树形图,所以模型一的对左右方向的定义在此需要改进。为此提出模型二:

从管道树的根部开始,以深度优先搜索找到一条管道线,并把它与原管道树分离,剩余部分为一个森林。再对其中的每一个非线形树做同样的处理,直到原管道树被分解为多条管道线。对于重复出现的节点,增加虚拟节点。如以 A_1 为根节点对两个管道树进行研究,A_{11} 节点有两个儿子,所以要增加一个虚拟节点 A'_{11} 与 A_{17} 相连(对 A_{17} 和 A_9 也要增加虚拟节点)。另外,我们规定从父亲到儿子的方向为右,反之为左。如 $A'_{11} \to A_{17}$ 为向右,$A_{17} \to A'_{11}$ 为向左。本题转化为对多条管道线的铺设优化问题,再用模型一中的方法对此问题求解。从 S_i 到 A_j 的最小运费表(虚拟节点与原节点的 $F_{i,j}$ 相同)和钢管订购和运输计划如表 5、表 6 所示:

表 5 最小运费表(问题三)

	S_1	S_2	S_3	S_4	S_5	S_6	S_7
A_1	170.7	215.7	230.7	260.7	255.7	260.7	275.7
A_2	160.3	205.3	220.3	250.3	245.3	250.3	265.3
A_3	140.2	190.2	200.2	235.2	225.2	235.2	245.2
A_4	98.6	171.6	181.6	216.6	206.6	216.6	226.6
A_5	38.0	111.0	121.0	156.0	146.0	156.0	166.0



续表 5

	S_1	S_2	S_3	S_4	S_5	S_6	S_7
A_6	20.5	95.5	105.5	140.5	130.5	140.5	150.5
A_7	3.1	86.0	96.0	131.0	121.0	128.1	141.0
A_8	21.2	71.2	86.2	116.2	111.2	116.2	131.2
A_9	64.2	114.2	48.2	84.2	79.2	84.2	99.2
A_{10}	92.0	142.0	82.0	62.0	57.0	61.0	77.0
A_{11}	96.0	146.0	86.0	51.0	33.0	47.0	64.0
A_{12}	106.0	156.0	96.0	61.0	51.0	37.0	56.0
A_{13}	121.2	171.2	111.2	76.2	71.2	16.2	38.2
A_{14}	128.0	178.0	118.0	83.0	73.0	11.0	26.0
A_{15}	142.0	192.0	132.0	97.0	87.0	28.0	2.0
A_{16}	60.0	110.0	44.0	80.0	75.0	80.0	95.0
A_{17}	95.0	145.0	85.0	50.0	32.0	46.0	63.0
A_{18}	100.0	150.0	90.0	55.0	45.0	33.0	50.0
A_{19}	105.0	155.0	95.0	60.0	50.0	36.0	55.0
A_{20}	115.0	165.0	105.0	70.0	65.0	10.0	32.0
A_{21}	125.0	175.0	115.0	80.0	75.0	0.0	26.0

表 6　钢管购运和铺设计划(问题三)

	S_1	S_2	S_3	S_4	S_5	S_6	S_7
$A_1 \to A_2$	0	0	0	0	0	0	0
$A_2 \to A_1$	0	104	0	0	0	0	0
$A_2 \to A_3$	0	75	0	0	0	0	0
$A_3 \to A_2$	0	0	226	0	0	0	0
$A_3 \to A_4$	0	223	59	0	0	0	0
$A_4 \to A_3$	0	98	0	0	370	0	0
$A_4 \to A_5$	0	0	0	0	0	0	0
$A_5 \to A_4$	334	0	0	0	272	0	0
$A_5 \to A_6$	0	0	9	0	0	0	0
$A_6 \to A_5$	185	0	0	0	0	0	0
$A_6 \to A_7$	16	0	0	0	0	0	0
$A_7 \to A_6$	189	0	0	0	0	0	0

续表 6

	S_1	S_2	S_3	S_4	S_5	S_6	S_7
$A_7 \to A_8$	76	0	0	0	0	0	0
$A_8 \to A_7$	0	125	0	0	0	0	0
$A_8 \to A_9$	0	175	0	0	0	0	0
$A_9 \to A_8$	0	0	505	0	0	0	0
$A_9 \to A_{10}$	0	0	159	0	0	0	0
$A_{10} \to A_9$	0	0	0	0	126	195	0
$A_{10} \to A_{11}$	0	0	0	0	0	30	0
$A_{11} \to A_{10}$	0	0	0	0	270	0	0
$A_{11} \to A_{12}$	0	0	0	0	110	0	0
$A_{12} \to A_{11}$	0	0	0	0	0	110	0
$A_{12} \to A_{13}$	0	0	0	0	0	1	0
$A_{13} \to A_{12}$	0	0	0	0	0	209	0
$A_{13} \to A_{14}$	0	0	0	0	0	184	0
$A_{14} \to A_{13}$	0	0	0	0	0	236	0
$A_{14} \to A_{15}$	0	0	0	0	0	335	0
$A_{15} \to A_{14}$	0	0	0	0	0	165	0
$A_{11} \to A_{17}$	0	0	0	0	0	0	0
$A_{17} \to A_{11}$	0	0	0	0	10	0	0
$A_{17} \to A_{19}$	0	0	0	0	95	0	0
$A_{19} \to A_{17}$	0	0	0	0	0	95	0
$A_{19} \to A_{20}$	0	0	0	0	0	0	0
$A_{20} \to A_{19}$	0	0	0	0	0	260	0
$A_{20} \to A_{21}$	0	0	0	0	0	0	0
$A_{21} \to A_{20}$	0	0	0	0	0	100	0
$A_{17} \to A_{18}$	0	0	0	0	50	0	0
$A_{18} \to A_{17}$	0	0	0	0	0	80	0
$A_9 \to A_{16}$	0	0	0	0	0	0	0
$A_{16} \to A_9$	0	0	42	0	0	0	0
T_i	800	800	1000	0	1303	2000	0

在上表的方案下，问题三的最小总费用为 140.66 亿元。

六、模型求解计算的优化

在模型一、二中,变量的个数随着 S_i, A_j 个数的增加而增加,造成搜索速度变慢。比如问题一中有 15 个 A_j,7 个 S_i,要铺设的管道有 14 段,由于每段管道都从与它关联的两个 A_j 铺设,要设的变量有 $14\times2\times7=196$个,而在问题三中变量数达到了 $20\times2\times7=280$ 个。但是有很多方面可以优化,以减少变量的数目:

(1)对于某个 A_j,如果从每个 S_i 到它的运费最小路径都经过同一条与它相关联的待铺设管道段,则铺设此管道段时必定不从 A_j 开始铺设。(否则先把钢管沿此管道段运到 A_j,再回头铺设,造成浪费)如问题一、三中的 A_1A_2 和 A_4A_5 路段的铺设方法只能是 $A_2\rightarrow A_1$ 和 $A_5\rightarrow A_4$。

(2)如果去掉钢厂产量上限的约束来求目标函数,搜索速度要快很多,虽然求得的结果不是原目标函数的解,但通过对它的分析能得到一些定性的结果,以减少变量和检验最优解。如对问题一,在 $S_i(i=1,2,\cdots,7)$的产量上限都没有限制的情况下,求得的最优解中从 S_1 运出的钢管最多,但是没有从 S_1 运送钢管到 $A_1,A_{10},A_{11},A_{12},A_{13},A_{14},A_{15}$这 7 个节点。也就是说,在除 S_1 外的 6 个厂满足以上 7 个节点钢管需求的条件下,从 S_1 运送到这些节点的钢管量为 0。对照问题一的结果,发现与此符合。再去掉供应钢管量最大的 S_1 厂,对剩余 6 个厂产量在不加限制的条件下求解,得到 S_2 供应钢管量在这 6 个厂中最大,但是没有从 S_2 运送钢管到 $A_1,A_9,A_{10},A_{11},A_{12},A_{13},A_{14},A_{15}$这 8 个节点。用同样方法我们得到一些在最优解中不太可能出现的运送途径,如表 7 所示:

表 7

	在最优解中不太可能出现的运送途径
S_1	$A_1,A_{10},A_{11},A_{12},A_{13},A_{14},A_{15}$
S_2	$A_1,A_9,A_{10},A_{11},A_{12},A_{13},A_{14},A_{15}$
S_3	$A_1,A_{10},A_{11},A_{12},A_{13},A_{14},A_{15}$
S_4	$A_i(i=1,2,\cdots,15)$
S_5	A_1,A_{13},A_{14},A_{15}
S_6	$A_i(i=1,2,\cdots,11)$
S_7	$A_i(i=1,2,\cdots,14)$



以上结果与问题一的结果非常吻合，所以在对问题一的目标函数求解时，如果去掉与之对应的变量，能使变量数目大大减少，加快搜索速度，且能得到最优解。但是对于钢厂产量上限约束较紧的情况，这个方法并不可行，得到的解可能不是最优解，但仍是一个较优的解。

七、模型的评价和改进方向

在实际的管道铺设问题中，一般要铺设的管道图是由多个树和线所组成的（反之如果有回路，就有不同路径到达某处，就需要更多的费用来购买、运输和铺设钢管，造成浪费）。模型二的优点在于它适用于由树、线组成的所有管道图。而缺点在于变量个数可能会很多，需要用到本文第六部分中的一些优化方法。

另外，注意到此问题中的非线性部分是一个二次型，其矩阵具有由全一矩阵组成的对角矩阵的固定形式，通过合同变换，将其化为形式很优的稀疏对角矩阵，可以使非线性部分有简单的形式，将约束条件重写后解不变，从而简化问题的处理。

八、参考文献

[1]叶其孝.大学生数学建模竞赛辅导教材(一).长沙:湖南教育出版社，1998

[2]张培强.MATLAB语言——演算纸式的科学工程计算语言.合肥:中国科技大学出版社，1995

[3]成礼智等.数学建模的理论与实践.北京:国防科技大学出版社,1999

[4]钱颂迪等.运筹学.北京:清华大学出版社,2000

点 评

本文对钢管订购与运输问题建立了一个数学规划模型。在建模中用最短路方法计算从钢厂 S_i 到枢纽点 A_j 的运费；针对每个钢厂要么不生产，要么至少生产 500 个单位这一约束，文章用类似于分支定界法的思路来处理。

文章对赛题提出的所有问题均给出了较理想的解决，最后还讨论了如何优化计算的问题。

由于此题属一类计算性质的赛题，题目要求明确，供选手发挥的余地相对较小，因此计算方法的正确性，结果的精确性和论文的表达就显得非常重要。本文作者在这些方面做得较突出，不仅很好地理解了题意，建立了合理的数学模型和有效的计算方法，结果准确，而且写作清晰流畅，言简意赅。

值得一提的是由于此题中不考虑铁路公路间转运的中转费用，也不限制转运次数，因此在计算从 S_i 到 A_j 的运费时许多参赛队往往考虑不周全。例如在计算从 S_7 到 A_{10} 的购运单价时，大多数参赛队都是先通过铁路（共 1130 公里，即 70 万元），再通过公路（70 公里，即 7 万元），从而得到单位运费为 77 万元的结果。但事实上，最少的单位运价应当通过以下方法得到：先通过长为 20 公里的公路从 S_7 到 A_{15}，接着从另一条等长的公路折返回铁路（共 40 公里，即 4 万元），然后经铁路（共 1130 公里，即 65 万元），再经公路（70 公里，即 7 万元）运到 A_{10}。这样，运价为 76 万元。不过这个问题对计算结果的影响不大，更不影响本题的解题思路。但值得注意的是，一些参赛队采用普通的运输问题模型先将各厂的钢管运到各枢纽点 A_j，再由 A_j 起向各方向铺设的“两阶段”建模方法，这是不妥当的，因为运输和铺设这两项费用是相互联系的，必须将他们综合起来考虑。另外，题目要求公路运价不足 1 公里按 1 公里计算，有些参赛队将公路运费看成公路里程的连续函数，于是目标函数的第 2 个和式中的 $L_j(L_j+1)/2+R_j(R_j+1)/2$ 变为 $\int_0^{L_j} x\,\mathrm{d}x = \int_0^{R_j} x\,\mathrm{d}x = (L_j^2+R_j^2)/2$，这也是不完全确切的。

本文获得 2000 年全国大学生数学建模竞赛全国一等奖。



血管的三维重建

金万军 赵立鸣 莫林剑

摘 要

本文对 A 题中给出的血管的三维重建问题进行了比较详细的讨论。围绕三维血管的中轴线的找寻和球半径的计算提出了一系列的解决办法,并逐步探求高准确度,低复杂度,鲁棒性好的数学模型。首先,我们基于对单个切片的分析,初步建立了“简单搜索”,“优化搜索”两个简单直观但时间复杂度较高的模型。然后,我们提出了血管全投影的概念,并基于此建立了三个风格各异的模型。模型 A 先求得中轴线在公共平面的投影,再回到每个切片计算出中轴线准确的空间位置和半径的长度。模型 B 着眼于单个切片和所有切片的联系,将一个切片的轮廓线与血管全投影的轮廓线联立,利用几何关系得到中轴线与此切片的交点位置及球半径的值。模型 C 最大可能地利用了单个切片上的信息,在每一个切片平面上除了计算中轴线与之的交点位置,对平面上其他中轴线投影线上的点也恢复出其深度值(Z 值),从而每一个(x,y)对应的 Z 值都是从所有包含它的投影的切片中计算出的一个个 Z 值综合而来。最后,我们对这三个基于血管全投影的模型进行了详细定量的比较和优缺点分析,并对题目中给定的三个假设约束进行了相关讨论。

一、问题的重述

1. 背景

医学序列的三维重建是近些年来医学图像和科学计算可视化领域热



门的研究课题。为了认识和研究人体或生物体的有关部分的内部结构和变化,人们利用的是分解和合成的办法,分解的过程是用一簇等间距平行平面将人体和生物体有关部分切成一片片薄片,即分解成切片的序列。通过对切片的观察,人们可以了解其内部的结构与变化。合成的过程就是将每一切片按顺序叠加起来,形成原有的空间结构,也就是从切片序列合成出原物的三维模型。这些切片可以是真切片,即生理切片,也可以是概念切片,如已经得到广泛应用的计算机断层扫描(CT)和核磁共振(MR)技术。合成过程可以由电脑进行,如基于计算机的曲面重建,也可以由人脑想象,在头脑中想象物体的空间结构。当然,使用电脑进行合成是趋势。

2. 问题

假设某些血管可视为一类特殊的管道,该管道的表面是由球心沿着某一曲线(称为中轴线)的球滚动而成的包络。例如圆柱就是这样一种管道,其中轴线为直线,由半径固定的球滚动包络形成。现有某管道的相继100张平行切片的图片(存成 BMP 格式的文件),记录了管道与切片的交点,要求根据这些切片来计算管道的中轴线与半径,给出具体的算法,并绘制中轴线在 XY、YZ、ZX 平面的投影图。

二、模型的条件和假设

1. 为简化模型而做的假设

(1)管道中轴线与每张切片有且只有一个交点;

(2)球半径固定;

(3)切片间距以及图像象素的尺寸均为 1。

如果不满足以上的任何一个假设,将有可能使本模型变得相当复杂,我们在第 5 部分将会对相关内容进行讨论。

2. 对输入图片的要求

题目所给的切片图像不是原始的血管切片的扫描图像,而是经过滤波,边界提取等预操作后的二值化 BMP 图片。在 BITMAP 的一位存取中,切片的实体区域用黑色像素(值为 0)表示,其余区域用白色像素(值



为 1)表示。

三、符号约定

Ω_i：序列号为 i 的切片（对应于第 i 张 BMP 图片）。简单起见，我们也用此符号代表切片内的黑像素区域；

Ω：公共平面，一个平行于 XY 坐标平面的平面。简单起见，我们也用此符号代表公共平面内的黑像素区域，称为血管全投影；

$\partial\Omega_i$：切片 i 上黑像素区域的边界；

$\dot{C}$：中轴线，由球心轨迹所组成的一条空间曲线段；

C：中轴线在公共平面上的投影曲线段；

C_i：中轴线在切片 Ω_i 中的投影曲线段；

O_i：$\dot{C}$ 与 Ω_i 的交点，即切片 i 上区域的最大内切圆；

R：生成球的半径；

投影：如没有特别指出，这里的投影都指的是正平行投影；

实体区域：三维物体在特定平面上的投影区域（如切片内的黑像素区域）；

截面圆：球和一个平行于 XY 坐标平面的平面相交所形成的圆；

投影圆：截面圆在切片上的投影；

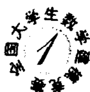
最大截面圆：过球心的截面圆；

最大投影圆：最大截面圆在切片上的投影。

四、问题的分析和解答

1. 概述

根据题意，血管的表面可看作由球心沿着某一曲线（称为中轴线）的球滚动包络而成。我们利用这一性质充分挖掘切片本身，单个切片与切片集合之间的性质关系，从多方面入手，定位中轴线与切片的交点和计算生成球的半径。实际上，由于切片的惟一最大内切圆的圆心便是中轴线与切片的交点，半径就是生成球的半径，所以中轴线和半径之间只要求出一个，就可以很容易地得到另一个。



2.基于单个切片 Ω_i 的分析计算模型

简单搜索模型

算法简述如下:

$$\forall S \in \Omega_i,\quad d_{S,\min} = \min(|PS|),\quad P \in \partial\Omega$$

当 S 跑遍 Ω_i, $r_i = \max(d_{S,\min})$ 此时对应的 S 即为 O_i。

对于每一个 $\Omega_i(i=0,1,2,...,99)$,根据以上算法得到与之对应的惟一的 $O_i(i=0,1,2,\cdots,99)$,连接起来即为管道的中轴线。

令所有 O_i 的 x 分量为0,即得到中轴线在 YZ 平面上的投影。

令所有 O_i 的 y 分量为0,即得到中轴线在 XZ 平面上的投影。

令 $R = \frac{1}{100}\sum_{r=0}^{99} r_i$,即得到球的半径优化搜索模型。

简单搜索需要对 Ω_i 中的左右点进行检验,时间复杂度大约为 10^8,所以我们从以下两个方面对简单搜索进行优化。

①利用半径的下界减小搜索区域

②利用估测半径进行匹配找点

但是由实际操作来看,即使这样,时间复杂度也在 10^8 左右,程序运行的速度不能满足对100张图片进行较快速度扫描的需要,不是一个经济的方法。所以,我们没有去深入研究此模型。

3.基于血管全投影的分析计算模型

二值BITMAP中,图片中黑色像素的像素值为0,白色像素的像素值为1。对100张切片图像进行逻辑操作(见图1)。因为所有切片两两平行,且切片图像关于 XY 坐标轴对齐,所以此操作相当于把所有的切片投影到一个公共平面 $Z=0$ 上。

实体变为

$$\Omega = \bigcup_{i=0}^{99} \Omega_i$$

区域 Ω 就是三维空间中的血管向二维公共平面上的投影。另外,由题目可知血管表面是由球心沿着某一曲线(称为中轴线)的球滚动包络而成,实质上就是无穷多个连续的球的包络面。所以,区域 Ω 的边界 $\partial\Omega_i$ 实际上也是连续的包络线,它包络了无穷多个由球向公共平面投影而构

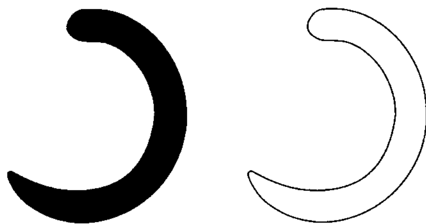


图 1

成的投影圆。

另外,为了以后分析的方便,我们通过 x 轴、y 轴的精化扫描线算法得到区域 Ω 的边界$\partial\Omega$,以及 Ω_i 的边界$\partial\Omega_i(i=0,1,\cdots,99)$,保存起来以备后用。

模型 A:在 Ω 中求 C, 然后在 Ω_i 中利用 C_i 恢复 O_i,从而得到$\dot{C}$。

$\dot{C}$是无穷多个球的球心连线,所以它在平面 Ω 上的投影 C 也是无穷多个投影圆的圆心连线,且这些投影圆都是区域 Ω 的相切圆。在本模型中,我们首先通过区域 Ω 计算出 C,再将 C 投影到切片 Ω_i 上得到 C_i,最后在 C_i 上求得 O_i 的位置。我们对每一个切片 $\Omega_i(i=0,1,2,\cdots,99)$ 求得其对应的 O_i,从而得到完整的$\dot{C}$。为了完整地得到中轴线在公共平面上的投影曲线段 C,有如下两种办法:

(1) 搜索求点法

定理 1 对于上的任意一点 P, 存在且仅存在一个以 P 为圆心,R 为半径的圆,使得此圆完全落在区域 Ω 中,且恰好与$\partial\Omega$ 相切。

由于空间中球的运动是连续的,所以 Ω 也是由连续的圆构成的。(形状图 2 所示)除了实体两端较尖的部分所包络的投影圆不是最大投影圆之外,实体的中间部分包络的都是最大投影圆,记为 O_i,他们具有相同的半径 r, 且 $r=R$。

定理 2 平面 Ω 上,任意最大投影圆 O_i 与区域边界$\partial\Omega$ 交于两点 P_{i1} 和 Q_{i1},这两点所在地方的曲率相等,$|P_{i1}Q_{i1}|$等于最大投影圆的直径。

证明:

如图 2 所示,圆 O_1 和 O_2 都是 Ω 上的最大投影圆。且圆 O_2 是由圆



O_1 向前移动距离 ΔL 得到的。由于 ΔL 足够小，所以在 Q_1 和 P_1 处的切线值可分别用 $\overrightarrow{Q_1Q_2}$ 和 $\overrightarrow{P_1P_2}$ 近似表示。因为圆 O_1 和圆 O_2 半径相等，且 ΔL 足够小，所以可知 $\overrightarrow{Q_1Q_2} \parallel \overrightarrow{P_1P_2}$ 且 $|\overrightarrow{Q_1Q_2}| = |\overrightarrow{P_1P_2}|$，即 $Q_1P_1P_2Q_2$ 构成一个长方形。Q_1，P_1 处的切线值相等。从而证得 Q_1 和 P_1 处的曲率相等。

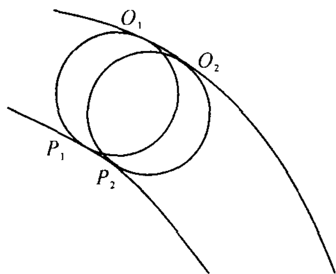


图 2

而又由于 $\overrightarrow{Q_1P_1}$ 与 Q_1 的法线方向重合，所以 Q_1P_1 是圆 Q_1 的一条直径，证毕。

推论： 如果在一侧弧上固定一点 Q，另一侧弧上找到一点 P，使得 $|PQ|$ 最短，则 Q 和 P 一定是一个最大投影圆的直径；P，Q 连线的中点为最大投影圆的圆心。

根据上面的推论，我们先在一侧弧上找一个点，然后扫描另一侧弧，找到一个点使得两点构成的线段长度最短。连接两点并取其中心，则此中心点一定在 C 上。我们将一侧弧上的点陆续试过之后，就可以得到所有 C 上的点了。我们同时可以得到一系列半径的值，经过求均值可以得到球半径的值 $\bar{R}$。后续模型的半径的值均取此值。实际情况表明，这样的选取以及较早定下 $\bar{R}$ 对以后模型的计算能起到了较好的效果，下面附一个半径表（表 1）。

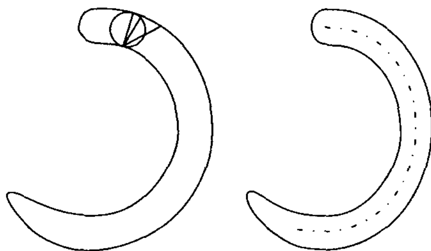

图 3



表1 半径表

R	29.50424	29.34706	30.23243	29.51694	29.34706
29.70690	29.50000	29.83287	29.70690	29.50000	30.22003
29.70690	29.50000	29.12044	29.69848	29.50424	29.73214
29.70690	29.50424	29.41088	29.30017	29.51694	29.83287
29.70290	29.50000	29.54657	29.96665	29.51694	29.70269
29.70290	29.50424	29.76995	29.82030	30.22003	29.74054
29.70290	29.50424	29.74054	29.52965	30.08322	30.04164
29.70290	29.70690	29.75315	29.52930	30.08322	30.22003
29.70290	29.70690	29.61841	29.52930	45.80393	29.90401
29.27456	29.53811	29.80772	29.56772	29.41088	29.88729
29.27456	29.03877	29.52965	29.26175	29.34706	29.60152
29.51694	29.27456	29.76995	29.84125	30.00417	30.18692
29.51694	29.74054	29.82030	29.61419	29.77415	29.80772
29.51694	29.51694	29.70269	29.88729	29.70269	29.60152
29.50424	29.76995	30.01666	29.76995	29.97082	29.47032
29.50424	29.83287	29.60152	29.83287	29.70690	29.60574
29.50424	29.34706	29.90819	30.37269	30.30264	
$\bar{R}=29.81285$					

由于中心点像素占据了一个像素，所以，实际用来搜索的 $r=\bar{R}-0.5=29.31285$。

(2) 曲线方程法

定义 参数曲线 $r(t)=(x(t),y(t))$，和它距离为 d 的参数曲线为：

$$r_d(t)=r(t)+d\cdot n(t)$$

$$n(t)=y'(t)-x'(t)/\sqrt{x'^2(t)+y'^2(t)}$$

我们称 $r_d(t)$为和 $r(t)$距离为 d 的等距曲线。其中 $r(t)$为母线，d 为带符号的偏移量，$n(t)$为单位法向量。

区域 W 两侧的曲线是无穷多个等半径的圆的包络线，由包络线的性质可知，一侧的曲线是和另一侧曲线距离为 $2R$ 的等距曲线，则 C 可被看作是跟两侧曲线距离为 R 的等距曲线。我们只要先用多项式拟合两侧曲线，就可以容易地得到 C 的函数表达式。

在本模型中，我们采取搜索求点法来得到 C，以保证更高的准确度。

得到了中轴线在公共平面上的投影曲线段 C 以后，根据垂直投影的不变性，我们把 C 投射回每一张切片 Ω_i，从而求出C与 Ω_i 的交点。

以第 20 张切片为例(见图 4),现将 C 投影到Ω_{20}上得到 C_{20},C_{20}与$\partial\Omega_{20}$相交于两点 P_{20},Q_{20}。然后进行如下操作:

$$\forall S\in P_{20}Q_{20},$$

$$d_{S,\min}=\min(|PS|),\quad P\in\partial\Omega$$

当 S 跑遍 $P_{20}Q_{20}$,$\max(d_{S,\min})$对应的 S 时,即为 O_i。

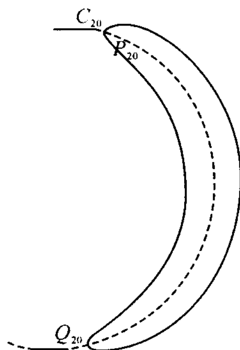


图 4

模型 B:通过 Ω 和Ω_i 恢复O_i,从而得到C

以上所提供的方法都需要先算出 C,而计算 C 所需要做的欧氏距离计算或者曲线拟合,都是比较费时的。

我们注意到,每一个切片都包络着大大小小连续的投影圆。但由于C和每一个切片 Ω_i 有且仅有一个交点,则每一个 Ω_i 有且仅有一个半径为R 的最大投影圆。这个最大投影圆 O_i 和$\partial\Omega_i$ 恰好相切于两点。所以我们只要将$\partial\Omega$ 和$\partial\Omega_i$ 取逻辑或非的关系,即可直接得到圆 O_i 和$\partial\Omega_i$ 相交的两点,很容易地求得 O_i 和半径R(见图 5)。

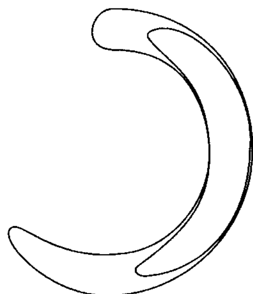


图 5

然而由于图像像素的近似性,$\partial\Omega$ 和$\partial\Omega_i$ 可能有不止一个点重合,因此我们设计了一个算法对此进行优化。

模型 C:C_i 基于深度恢复的分析计算:

在上面的模型讨论中,我们只是利用了其中最大内切圆的圆心和半径,而没有利用其他内切圆的信息,造成了很大的信息浪费。本算法改进了这个不足,具体算法如下:

①求等距曲线(见算法 A),得到一系列有序坐标存入矢量 x_c,y_c,其中$(x_c(i),y_c(i))$表示等距曲线上第 i 个像素的坐标。

②等距曲线是圆的内切圆心的轨迹。对 100 幅切片图 (Ω_1,…,Ω_{100})分别沿等距曲线逐个像素遍历,求得每个对应内切圆的半径。这组内切圆半径存放于矩阵 $K(l,i)$中,满足 $K(l,i)$为在 Ω_l 上以$(x_c(i),y_c$



(i))为圆心的内切圆半径,如果$(x(i),y(i))\notin \Omega_i$ 则$K(l,i)=0$

③由于血管可看成是由球体滚动形成,而球体的横截面是圆 $r^2+(z-z_0)^2=R^2$,其中 z 为当前截面圆圆心的纵坐标,z_0 为球心的纵坐标,R 为球半径。所以当 R、z_0 固定时,(r,z)的轨迹是半径为 R 的圆,并且这个轨迹圆的圆心就是球心。而当 i 固定时,由②可得 $r=K(l,i)$,$z=l$,所以$(K(l,i),l)$在以球心为圆心、半径为 R 的圆上。两个$(K(l,i),l)$的中位线和纵轴的交点可以确定 z_0。$(K(l,i),l)$和$(K(l,i+j),l)$构成一对数据,其中 j 从 5 递增到 15,对每一个 j,i 顺次取遍所有适用值。最后对求得的所有 z_0 求均值作为球心的纵坐标。

④如上,可得中轴线上点的纵坐标,结合等距曲线,可恢复中轴线上一系列点的三维坐标。

优点:

a.比较前面的算法,本算法充分利用了题目给出的数据,故结果具有更好的精确度。

b.通过(x,y)得到 z,所以得到的中轴线上点的坐标更多(数量约 5:1)。

c.适应性较好,应用面比较广泛。

缺点:

a.得到的数据在两个端点处不甚理想,原因为:

- 等距曲线的两端没有严格限定;
- z 的精度依赖于 $K(l,i)$的非零数据的多少,而在两端处这样的数据显然比中间少;
- 两端的尺寸小,BMP 格式带来的误差影响增大。

b.对于得到的所有曲线,由于是取点产生的,所以不连续,而 MatLab 自动以直线连接相邻点,所以造成中轴线的折拐。

现实中,血管不可能出现突然的折拐,所以我们采用 MatLab 提供的 Polyfit 函数对所得的曲线进行多项式拟合。多项式拟合的次数越高,生成的曲线越多地通过给定点,但是如果次数过高,则生成的曲线抖动严重。实验表明,4～10 次的多项式拟合效果较好,为了表达简便,我们采用 8 次多项式拟合。

对拟合采用的方式,我们采取如下办法:在 XZ 平面上拟合,使 $X=f(Z)$,在 YZ 平面上拟合,使 $Y=f(Z)$。由于 Z 为切片层值,固定

可靠,所以,这样的拟合将调整 X、Y,使得结果接近真实值。

拟合结果如下:

模型 A: $x=-0.0009z^4+0.0216z^3-0.2477z^2+1.1336z+94.9482$

$y=-0.0008z^4+0.0163z^3-0.1313z^2+0.3520z+256.9605$

模型 B: $x=0.0016z^4-0.0437z^3-0.5676z^2-2.9058z+99.0302$

$y=0.0007z^4-0.0210z^3+0.3349z^2-1.6910z+253.9568$

模型 C: $x=-0.0004z^4+0.0109z^3-0.1386z^2+0.7284z-161.8042$

$y=-0.0003z^4+0.0090z^3-0.0958z^2+0.4443z-0.3405$

其中,模型 C 所得的中轴线的三维图如下:

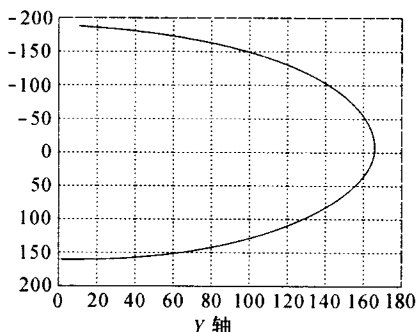


图6 XY 面投影

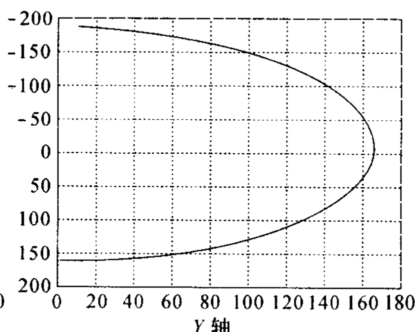


图7 XY 面投影

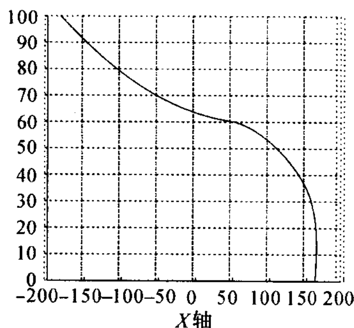


图8 ZX 面投影

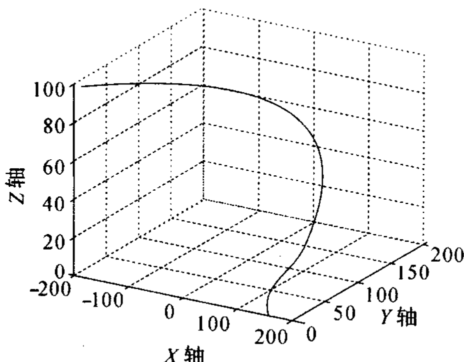


图9 中轴线的三维图

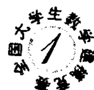


表 2　x,y,z 坐标表

Z	X	Y
0	−161	0
1	−161	0
2	−161	0
3	−161	0
4	−161	0
5	−161	0
6	−161	0
7	−161	1
8	−161	1
9	−161	1
10	−161	1
11	−161	2
12	−161	2
13	−161	2
14	−161	3
15	−161	4
16	−161	3
17	−161	5
18	−161	6
19	−161	7
20	−161	7
21	−161	9
22	−161	10
23	−161	12
24	−161	14
25	−161	16
26	−161	19
27	−161	21
28	−161	24
29	−160	28
30	−160	29
31	−160	32
32	−159	36
33	−158	38
34	−157	43
35	−156	47
36	−156	49
37	−155	52
38	−154	56
39	−152	60
40	−150	64
41	−149	68
42	−146	73
43	−144	77
44	−142	81
45	−138	87
46	−136	91
47	−132	96
48	−127	103
49	−125	105
50	−121	110
51	−116	115
52	−112	119
53	−107	123
54	−101	128
55	−95	133
56	−89	137
57	−83	141
58	−78	143
59	−68	149
60	−64	150
61	−53	155
62	−47	157
63	−39	159
64	−31	161
65	−22	163
66	−14	164
67	−5	165
68	6	166
69	15	166
70	20	165
71	31	164
72	38	163
73	50	161
74	60	158
75	67	156
76	75	153
77	86	149
78	94	145
79	101	141
80	109	136
81	115	132
82	126	124
83	129	121
84	137	114
85	146	104
86	149	101
87	155	93
88	160	86
89	164	80
90	168	73
91	171	67
92	175	58
93	177	53
94	180	45
95	183	36
96	184	31
97	186	23
98	187	17
99	188	10

表 3　三种模型的分析和比较

模型名称	优　　点	缺　　点
模型 A	由于使用了焦点求圆心的方法，各切面和中心轴的焦点计算比较准确。但是其稳定性低，半径的变化比较大	计算量较大，各切面和中心轴的焦点之间的点的定位不准确，使得最终的三维曲线不光滑，有较多折点
模型 B	对中心轴和切面的焦点的定位十分简单。计算量小。稳定性较高，半径的变化较小	对焦点的定位不如 A，C 模型准确，造成恢复血管切面时与原始图片有一定的出入
模型 C	在准确性、连续性和稳定性上表现最好。充分反映了三维血管模型的真实情况	计算量最大

判定算法的有效性

我们评价一个模型是否有效主要从下面 3 个方面考虑：

(1)准确性

因为我们的任务就是尽可能精确地恢复出三维空间中的中轴线，因而准确性是我们首要考虑的。我们采取了两种准则判定。

在原切片上进行检验：

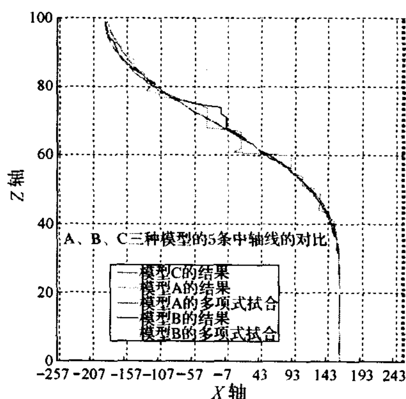


图 10

模型计算出来的中轴线与每个 Ω_i 相交于 O_i。以 O_i 为圆心，$\bar{R}$ 为半径做判定圆。令 outcount 为判定圆与 Ω 相切的点的个数。我们认为，交切(图 11)和邻切(图 12)都可以看作相切，交切点的集合为 S，邻切点的集合为 T。则 tancount = max$\{N(S), N(T)\}$。显然，我们希望 outcount 非常小，而 tancount 比较大。如果 outcount 小于阈值(本题取 4)，而 tancount＞阈值(本题取 10)，则称模型对这个切片有效，否则称模型对这个切片无效。

(2)鲁棒性

此模型所计算出的 R 的值序列，应该保持一个尽可能小的方差

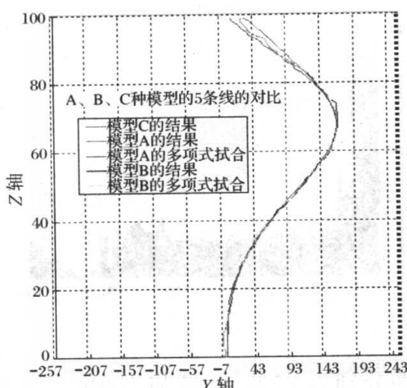


图 11

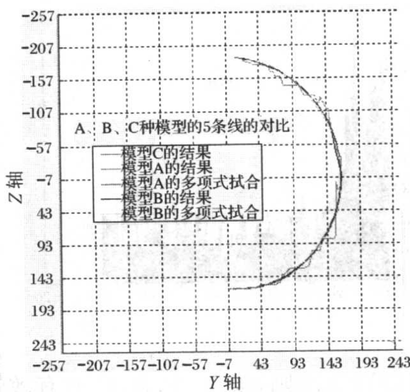


图12

$$\delta(R) = \sum (R_i - \bar{R})^2$$

(3)复杂性

此模型的计算复杂度应在可以接受的范围内(小于 10^{10})。

我们列出利用已得的中心轴曲线和切面焦点以及半径恢复内切圆图样并与原始图样检验的方法来对三种模型进行比较。下列组图中,红色柱状图表示恢复的内切圆与原始图样相切的像素个数(即 tancount),蓝色柱状图表示恢复的内切圆超出原始图样的像素个数(即 outcount)。(注:原图为彩色图)

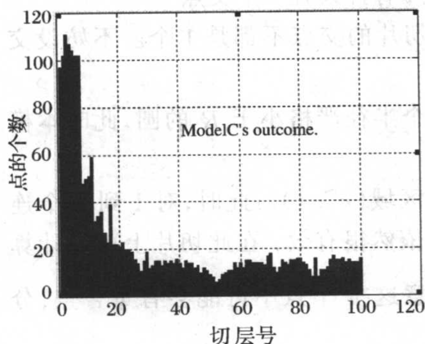


图 13

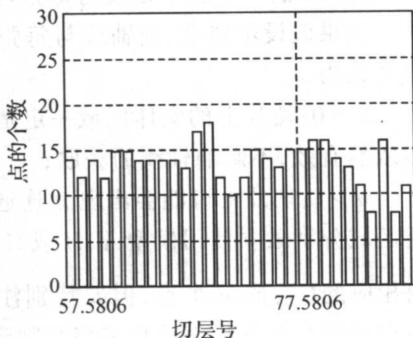


图 14

根据上面的判定,我们看到三个模型对本问题都十分有效。按照我

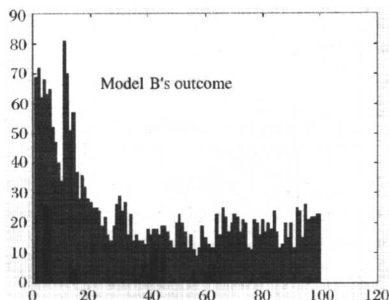


图 15

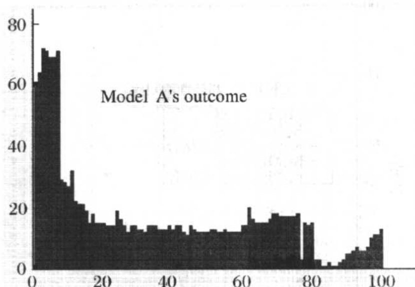


图 16

们的判定准则，模型 A，B，C 的有效性分别为 92%，90% 和 100%。可见模型 C 的效果最好。

五、进一步研究的问题

1. 进一步放松假设的约束

为了简化模型，题目中给出了三个假设。然而显然这些假设有些过于严格，实际情况很少符合这样的简化模型。我们以后的工作就是要不断放松这些假设的约束，建立更加鲁棒的模型。现在，我们来分析关于这个模型的三个假设，约定讨论一个假设时，其他假设仍满足。

[假设 a] 管道中轴线与每张切片有且只有一个交点

如果假设不成立，则轴线与每张切片的交点不都是 1 个。不妨设交点个数为 n。

$n=0$：切片上的实体区域一定是个半径严格小于 R 的圆，此时本模型依然有效，不影响轴线的获取；

$n>1$：切片上可能会出现 k 连通区域（$k\geqslant 1$）。此时，对 1 到 k 个连通区域分别独立使用模型 A，B 或 C，依然很有效，在此切片上可以计算出相应的 C 上的 n 个点，但要特别注意这 n 个点中可能会有重合点，分析中要结合上下相邻切片来综合判定

[假设 b] 切片间距以及图像象素的尺寸均为 1

如果切片间距大于 1，则所有的切片中的实体只是体数据的一个真子集。对于单轮廓线之间的三维形体重构问题，如果切平面之间间距小



且相互平行,可以先用最小生成树算法解决轮廓线对应问题,再用三角面片重构三维形体。如果在相邻的两层平面上,或者在其中的一层上,有多条轮廓线,则需要解决多轮廓线之间的三维形体重构问题。

第三种算法可以用来解决这种由部分切片恢复全部三维样本的问题。事实上,如果血管半径固定,我们只要有两幅图在直径范围内,就可以恢复到这两幅图中间的所有的三维图形。

首先,计算出它们的等距曲线,得到的最大内切圆半径也就是球半径 R。看两幅图是否在一个直径范围内,如果是,按照算法 C 得到等距曲线上的所有的内切圆的半径。球的(x, y)固定,两个数据,可以恢复它的 Z, R。对范围内等距曲线上的点都可以得到对应的中轴线上点的坐标。有了中轴线和球半径,可以恢复出三维图形。但是由于对取得的内切圆半径的精度要求很高,题目提供的 BMP 图片格式达不到要求,所以还要借助大量的切片数据来恢复,以提高精度。至于由部分切片恢复全部三维形状的问题,比如应该选取怎样分布的切片序列,这些切片要满足什么条件等等,以后会有很多工作要做。

[**假设 c**] 球半径固定

如果球半径不固定,半径 r 是空间坐标的函数 $r = f(x, y, z)$,那么问题会变得比较复杂。由于时间所限,我们没有进一步研究。

2. 对原始切片图像的处理

本题给出的切片图像是经过处理以后的切片。在实际的医学应用中,我们得到的切片图像(也叫断层图像)往往会有噪音大、分辨率低、边界模糊不光滑等特点,我们需要用图形图像处理的手段对其进行处理,如滤波,边缘检测,边界点提取和边界点数据插值与光顺等操作,从而得到可供我们模型分析解决的输入图像。

3. 基于空间体数据的绘制

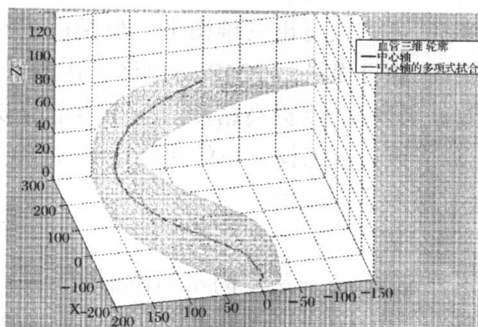
医学图像数据是三维正交的等距网格,其网格单元可看成是正六面体单元,整个数据场就是由这种正六面体组成的,这种组成三维图像的基本正六面体单元被称为体素(voxel)。在这里我们依据 Marching Cube 算法,将体素定义为相邻层之间八个网格点组成的数据单元。这与我们得

到的切片序列中的数据单元恰好是吻合的。因为我们所得的切片中的数据是以一个像素为单位,且相邻两层的距离也是一个像素,所以我们可以把切片序列中的数据单元看作三维空间中的体素,所有输入数据即为三维空间中的体素集,血管表面即为体素集的包络面。

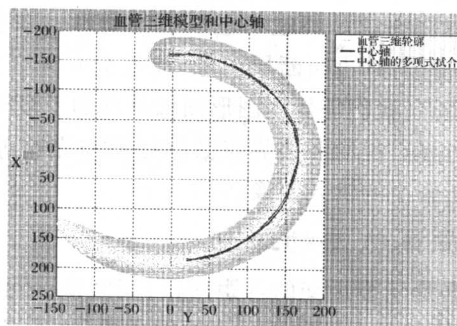
六、参考文献

[1]MATLAB 5 手册.北京:机械工业出版社
[2]计算机辅助几何设计(CAGD).浙江大学数学系
[3]医学图像处理.北京:高等教育出版社

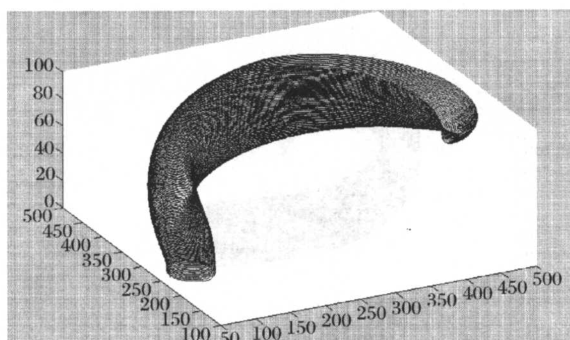
附录 1



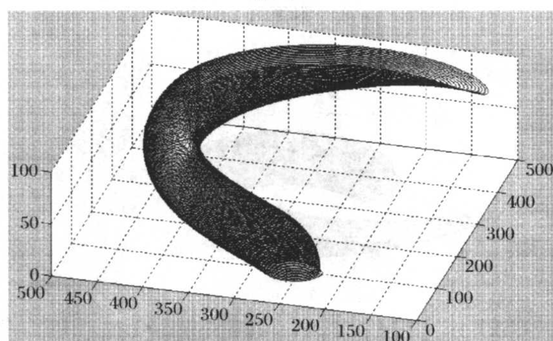
附图 1



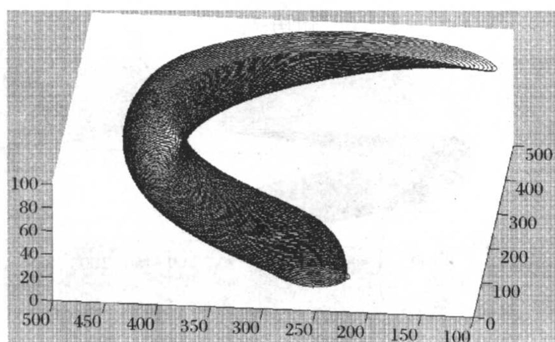
附图 2


100
80
60
40
20
0
500
450
400
350
300
250
200
150
100
50
100
150
200
250
300
350
400
450
500

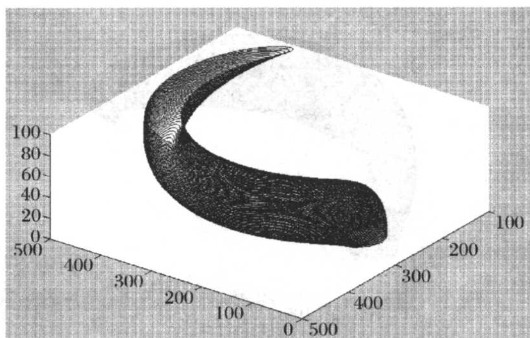
附图 3


100
50
0
500
450
400
350
300
250
200
150
100
0
100
200
300
400
500

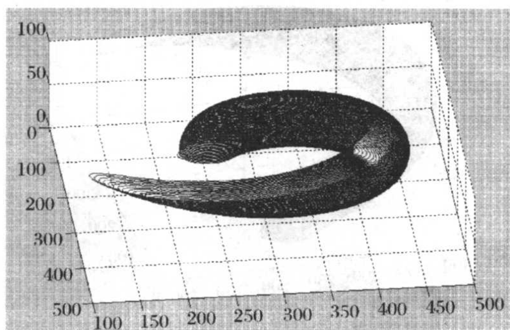
附图 4


100
80
60
40
20
0
500
450
400
350
300
250
200
150
100
0
100
200
300
400
500

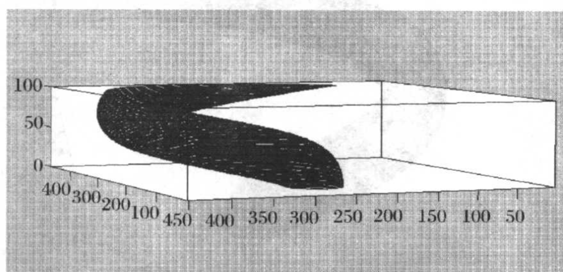
附图 5


100
80
60
40
20
0
500
400
300
200
100
0
500
400
300
200
100

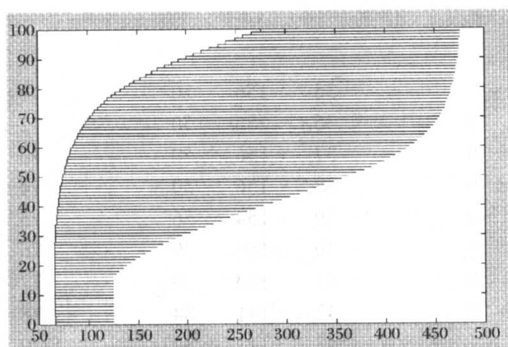
附图 6


100
50
0
0
100
200
300
400
500
100
150
200
250
300
350
400
450
500

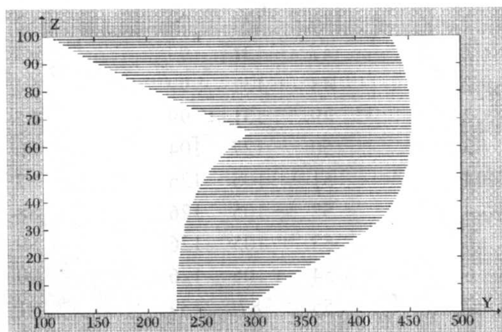
附图 7


100
50
0
400
300
200
100
450
400
350
300
250
200
150
100
50

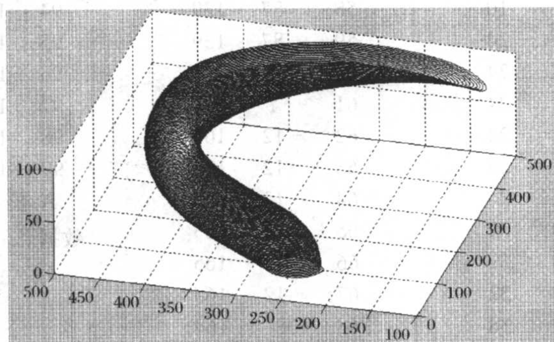
附图 8



附图 9



附图 10



附图 11



附录 2

Z	X	Y
0	−161	0
1	−161	0
2	−161	0
3	−161	0
4	−161	0
5	−161	0
6	−161	0
7	−161	1
8	−161	1
9	−161	1
10	−161	1
11	−161	2
12	−161	2
13	−161	2
14	−161	3
15	−161	4
16	−161	4
17	−161	5
18	−161	6
19	−161	7
20	−161	8
21	−161	9
22	−161	10
23	−161	24
24	−161	24
25	−161	24
26	−161	24
27	−161	24
28	−161	24
29	−160	32
30	−160	32
31	−160	32
32	−159	38
33	−159	38
34	−156	50
35	−156	50
36	−156	50
37	−156	50
38	−156	50
39	−154	56
40	−153	59
41	−153	59
42	−141	84
43	−141	84
44	−131	99
45	−131	99
46	−131	99
47	−131	99
48	−131	99
49	−131	99
50	−127	104
51	−105	126
52	−105	126
53	−105	126
54	−105	126
55	−87	139
56	−87	139
57	−87	139
58	−87	139
59	−87	139
60	−82	142
61	−1	166
62	−12	165
63	−12	165
64	−12	165
65	−12	165
66	−12	165
67	−12	165
68	38	164
69	38	164
70	38	164
71	38	164
72	38	164
73	38	164
74	38	164
75	53	161
76	75	153
77	109	137
78	109	137
79	128	123
80	119	129
81	119	129
82	130	120
83	130	120
84	130	120
85	146	104
86	146	104
87	146	104
88	146	104
89	160	86
90	160	86
91	166	76
92	166	76
93	175	58
94	177	53
95	177	53
96	180	45
97	181	42
98	183	35
99	184	31

A 模型的结果数据。



Z	X	Y
0	−161	−3
1	−161	−3
2	−161	−3
3	−161	−7
4	−161	−7
5	−161	−7
6	−161	−7
7	−161	−7
8	−161	−3
9	−161	−3
10	−161	1
11	−161	1
12	−161	2
13	−161	4
14	−161	5
15	−161	6
16	−161	7
17	−161	8
18	−161	9
19	−161	10
20	−161	10
21	−161	11
22	−161	12
23	−161	15
24	−161	16
25	−161	19
26	−161	20
27	−161	22
28	−160	24
29	−160	32
30	−159	33
31	−159	36
32	−158	41
33	−158	41
34	−157	44
35	−156	48
36	−155	52
37	−153	57
38	−151	63
39	−149	66
40	−148	70
41	−146	72
42	−145	75
43	−144	78
44	−139	87
45	−134	93
46	−133	96
47	−127	103
48	−122	108
49	−122	109
50	−115	116
51	−110	121
52	−108	123
53	−104	126
54	−95	133
55	−91	136
56	−87	138
57	−80	142
58	−74	145
59	−56	153
60	−50	156
61	−43	158
62	−39	159
63	−32	161
64	−22	163
65	−19	163
66	−9	165
67	0	166
68	9	166
69	10	166
70	10	166
71	10	166
72	18	165
73	17	165
74	17	165
75	51	157
76	78	152
77	90	147
78	96	144
79	110	136
80	112	134
81	115	132
82	124	126
83	128	122
84	138	113
85	147	103
86	151	98
87	157	90
88	163	81
89	166	77
90	170	70
91	172	65
92	174	61
93	175	59
94	180	45
95	183	35
96	184	32
97	185	27
98	187	20
99	186	23

B 模型的结果数据。

点 评

本文提出了血管三维重建的三种方法,即文章所说的模型 A、B、C。作者利用题目提供的数据,分别用这三种方法实现了血管重建,并对用这些方法重建的优缺点进行了细致的对比。

从论文中可以看出,作者在研究过程中总体设想明确、思路开阔,具有较强的研究能力。作者能沿着几条不同的途径去设计算法,并通过对比指出各种算法的优缺点,这种研究方法是值得提倡的。作者十分清楚,由于模型误差和计算误差的缘故,重建出来的血管与真实血管会有一定的差异,他们能将计算结果与原始切片数据进行一一对照,以此来评估算法的实际效果,对模型进行必要的检验,这种做法应予肯定。在文章的第五部分,作者提出了一些进一步研究的建议,这些建议具有较强的实际背景,显然具有进一步研究的价值,这样做也是值得鼓励的。

文章尚有一些不足之处。其一,作者在介绍算法时讲述得还不够清楚,阅读时有些地方还让人感到费解。文章是写给别人看的,在撰写时应尽量注意如何表达才能使别人易于理解,对于这一点,本文作者似乎还不够注意。其次,文章虽然总体思路较清楚,且可以看出他们已作了较好的研究工作,但文章本身却似乎未经仔细的修改,写得较为粗糙。科研论文是研究工作的结果,交稿前应认真推敲、反复修改,这也是一项不可轻视的工作。

本文获2001年全国大学生数学建模竞赛全国一等奖。



公交车的调度

杜克勤　童　颜　李　科

摘　要

本文解决的是一个公交车调度问题,目的是用尽可能少的车来运送乘客,同时不能让乘客等待时间过长,也不能超载。

文中提出了一种解决本问题的新颖的模型。通过分析知,一个时区内需要的车只与该时区内车站的最大转移客流量有关,于是我们对题目所给的数据进行一系列变换处理,求出每个时段发的车的最大转移客流量,得到一组新的数据。根据最大转移客流量,用线性规划的方法,可以求出我们这个模型需要发车次数的最小值,然后综合考虑乘客的等待时间等,我们可以给出各个时区的发车次数,进而确定全天发车时间表,由发车时间表,我们同样用线性规划方法求出需要的最少车辆数。

用我们构造的模型,求出了一个可行的调度计划,并给出了发车时刻表。每个方向的全天发车次数为237,需要57辆公交车。通过编程模拟得出平均等待时间为2.17分钟和平均满载率81.6%,对这个解进行了评价,说明了本模型的特点,并指出了进一步优化改进的方向。最后用一个模拟搜索方法又求出一个解,并与第一个解作了比较。

一、问题的提出

这是一个公共汽车调度问题。题目给出了来自我国一座特大城市某条公交线路的客流调查和运营资料。

该条公交线路上行方向共14站,下行方向共13站,题目给出的是典

型的一个工作日两个运行方向各站上下车的乘客数量统计。公交公司配给该线路同一型号的大客车,每辆标准载客 100 人,据统计客车在该线路上运行的平均速度为 20 公里/小时。运营调度要求,乘客候车时间一般不要超过 10 分钟,早高峰时一般不要超过 5 分钟,车辆满载率不应超过 120%,一般也不要低于 50%。

我们的目标就是根据题目所给的这些统计资料,把调度问题抽象成一个明确完整的数学模型,并求解,根据我们的解,给公交公司制定了一个公交车调度方案和起点站的发车时刻表,使公交公司能够有效降低成本,但又不能牺牲乘客的利益。

二、基本假设

(1)候车队伍有良好的秩序;即要保证乘客先来先乘车的原则;

(2)忽略其他情况对公交车的影响,即公交车以 20 公里/小时的速度匀速行驶;不计乘客上、下车的时间及公交车起动、加速、滑行、制动时间(因为题中给的是平均速度);

(3)从 i:00—($i+1$):00 时间段内到达某站的乘客数服从均匀分布;

(4)公交公司只对公交车进行调度,但在允许的范围内不限制乘客上车,即只要该车乘客数不大于 120 则允许乘客上车,直到达到 120 为止;

(5)每个时间段的发车间隔时间是确定的,而且是相等的。

三、符号说明与概念引进

1. 概念引进

(1)时区——我们规定两相邻正点之间的单位时间间隔为一个时区,并予以编号

5:00 - 6:00　　1 时区

6:00 - 7:00　　2 时区

⋮

21:00 - 22:00　　17 时区

22:00 - 23:00　　18 时区

(2)跨时区——当一辆公交车从 i 时区出发在到达目的地之前已经



进入下一个 $i+1$ 时区时，我们就认为产生了跨时区。

(3)站间转移客流量 $M(i,j)$——有多少乘客乘坐 i 时区发出的车经过 j 车站到达 $j+1$ 车站(包括在 j 车站上车的乘客)。

(4)如果一个乘客是在 i 时区来到 j 车站乘车的，我们称该乘客来自 i 时区

2.符号说明

(1)$l_1,l_2,l_3,\cdots,l_{18}$分别代表各个时区内发车的时间间隔

(2)$n_1,n_2,n_3,\cdots,n_{16},n_{17},n_{18}$分别代表各个时区内发车的总次数

(3)$A(i,j)$:在 i 时区 j 车站的净上车客流量

(4)$M(i,j)$:站间转移客流量(见上面定义)

(5)$T_{i-1}(i,j)$:j 车站来自 i 时区的人被发自 $i-1$ 时区的车运走的数量

(6)$N(i,j)$:i 时区内发出的所有车在 j 车站运走的净乘客数量(上车人数减下车人数)。

四、问题的分析

我们应该在满足乘客的等待时间要求的前提下，合适地安排发车次数，得到发车时刻表，使得发车次数尽可能地少，而且用到的车辆尽可能少。

我们的目标是求每一个时区内发车的次数，发车的次数为在该时区内发出的车数，取决于要运送的乘客数量，那么我们就要求出各个时区发出的车分别要运送的乘客数。知道了要运送的乘客数，就可以用线性规划求发车车次。知道了发车车次，进而可以列出发车时刻表，通过时刻表，可以用线性规划求出最少车辆数。

五、模型的建立和求解

1.模型一

示意图见图1。箭头表示车在一定时间内从起点开到终点。由于公交车有跨时间段的问题(例:起点到第 Ai 站要用 t 小时)的时间，则在 i:

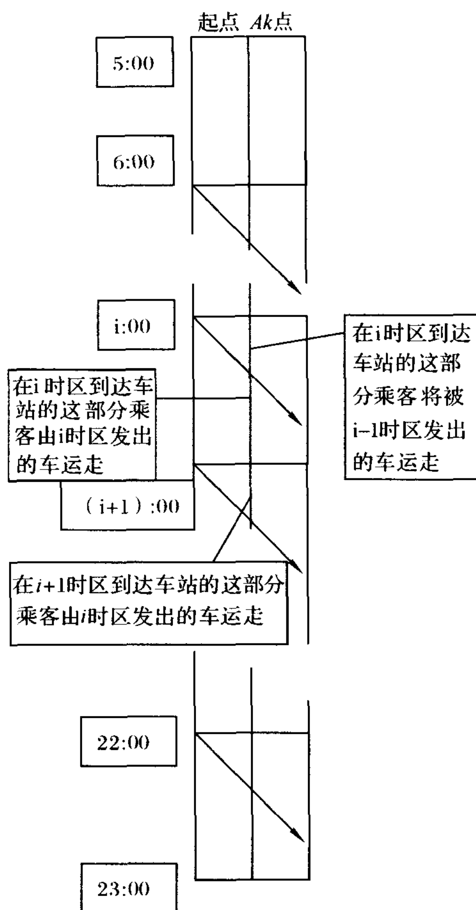


图 1

00 到 $i+t$ 时间内到达 Ak 站的乘客可由 $i-1$:00 到 i:00 时间段内发出的公交车运走)。为了方便解决问题,我们以第 i 时区($6\leqslant i\leqslant 22$)发出的公交车来重新划分乘客,找出各个站点上由这些车运送的乘客数量,则从第 j 时区发出的车在 Ak 站点所要运走的乘客数量为(假设在 i:00 到($i+1$):00 在 Ak 点候车的乘客数量为 u_i;在 i:00 到($i+1$):00 在 Ai 点下车的乘客数量为 d_i):



当 $j=1$ 时，上车的乘客数为：$u_5+u_6\times t$

下车的乘客数为：$d_5+d_6\times t$

当 $2\leqslant j\leqslant 17$ 时，上车的乘客数为：$u_i-u_i\times t+u_{i+1}\times t$

下车的乘客数为：$d_i-d_i\times t+d_{i+1}\times t$

当 $j=18$ 时，上车的乘客数为：$u_{22}-u_{22}\times t$

下车的乘客数为：$d_{22}-d_{22}\times t$。

我们再引入两个附加假设：

附加假设一：假设 i 时区发出的第一辆与最后一辆车到达 j 车站的时刻分别为 $t_1(j)$，$t_2(j)$，在 $t_1(j)$ 到 $t_2(j)$ 这个时间间隔内，乘客的到来服从均匀分布。

(原来假设一个时区里来人是均匀的；当求出在 $t_1(j)$ 到 $t_2(j)$ 内来的人数后，再认为这段时间内来人均匀)

附加假设二：假设 i 时区发出的车能够把 $t_1(j)-t_2(j)$ 这个时间间隔内来到 j 车站的乘客全部运走。

下面我们要解决的问题：具体求出 i 时区发出的车要运多少乘客。

命题一　在考虑车站转移客流量 $M(i,j)$ 时，在一个站我们分别考虑上下车人数与我们考虑净上车人数是等价的。

命题二　每个时区内公交车的发车次数的下限由这个时区内相邻两站间最大转移客流量 $\max M(i,j)$ 决定。

证明　$\max M(i,j)$ 表示在 i 时区要经由 j 车站向下的乘客，我们要保证人能被运走，但又不能超载，即满载率不能超过 120%，那么如果 $n_i<\dfrac{\max M(i,j)}{120}$，必然导致该站的部分乘客滞留。

命题三　在第 i 时区来 j 车站的乘客是被发自 $i-1$ 时区和发自 i 时区的车共同运走的，而且被 $i-1$ 时区的车运走的乘客 $T_{i-1}(i,j)$ 占 j 车站总乘客的比例由 j 车站距始发站的距离决定。

推论　第 i 时区内发的车从 j 车站运走的乘客来自于 i 时区和 $i+1$ 时区，运送的 $i+1$ 时区的乘客 $T_i(i+1,j)$ 占总乘客的比例也由 j 车站距始发站的距离决定。

由这四个命题我们可以得出的一个算法：



①计算 i 时区 j 车站净上车人数 $A(i,j)(i=1,2,\cdots 17,18;j=13,12,11,\cdots,2,1,0)$；

②计算 $T_{i-1}(i,j)$ 和 $T_i(i+1,j)$（$T_{i-1}(i,j)$ 为在 i 时区来的乘客被 $i-1$ 时区发出的车运走的数量，$T_i(i+1,j)$ 为从 i 时区发的车运送的在 $i+1$ 时区来的乘客）；

③计算 i 时区发出的车需从 j 车站运走的乘客的数量 $N(i,j)$，$N(i,j)=A(i,j)-T_{i-1}(i,j)+T_i(i+1,j)$；

④ 计算 i 时区 j 车站的转移客流量 $M(i,j)$，$M(i,j)=\sum_{i=13}^{j}N(i,j)$（对于上行方向来说），即 i 时区内发出的车从起点站到 j 站的净上车人数之和；

⑤据 i 时区内各车站的最大转移客流量 $\max M(i,j)$ 和需要的发车次数 n_i 的关系：

$$\max M(i,j)\leqslant n_i * 120$$

从而，确定 n_i 的最小值。

可用以上算法，将原表格变换成为一个新的表格，用来求发车次数。

我们编了一个程序来计算每个车站由 i 时区出发的车的转移客流量，最后求得的结果列表如下：（各个时区的最大转移客流量用粗体标出）

表 1　上行方向的各个车站的转移客流量

站名	A13	A12	A11	A10	A9	A8	A7	A6	A5	A4	A3	A2	A1
1	371	445	512	556	679	**723**	645	563	670	660	685	718	44
2	1990	2279	2514	2620	**2996**	2944	2594	2106	2209	2074	2020	1908	409
3	3626	4038	4326	4485	**4912**	4753	4259	3589	3638	3498	3422	3245	1965
4	2064	2273	2445	2505	**2667**	2577	2267	1903	1911	1841	1789	1682	845
5	1186	1308	1396	1423	**1519**	1422	1210	982	1002	960	925	864	289
6	923	1024	1093	1123	**1211**	1134	975	774	819	780	755	712	208
7	957	1081	1180	1227	**1346**	1289	1143	981	1041	995	968	932	458
8	873	969	1057	1090	**1190**	1138	1022	899	945	915	899	874	483
9	779	878	941	958	**1042**	1013	922	804	838	812	796	767	405
10	625	695	763	796	**876**	861	765	611	637	605	583	555	213
11	635	734	803	834	**924**	890	733	509	548	513	476	432	-195

续表

站名	A13	A12	A11	A10	A9	A8	A7	A6	A5	A4	A3	A2	A1
12	1493	1716	1875	1938	**2148**	2052	1805	1470	1557	1476	432	1335	292
13	2011	2265	2444	2495	**2700**	2527	2305	2020	2099	2031	1991	1922	1289
14	691	767	823	833	**886**	826	726	575	599	572	547	520	264
15	350	391	423	435	**464**	440	388	299	302	287	276	263	99
16	304	337	362	373	**405**	394	354	282	277	264	254	241	111
17	209	230	246	250	**265**	246	219	167	159	149	141	131	734
18	**19**	19	19	17	14	5	-4	-16	-19	-20	-24	-27	-36

由以上数据,我们可以建立一规划模型来求解 min n_i

目标函数:min n_i

s.t.: $\frac{\max M(i,j)}{n_i} \leqslant 120(j=1,2,3,\cdots 17,18)$

经过求解,对所得数据进行近似及权衡乘客等待时间,我们得到了以下一组时区车次安排,见表 2。

表 2　上行方向的各个时区的车次安排

时区	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
车次	6	24	41	23	13	11	12	10	9	8	8	18	23	8	6	6	6	4

可见,总共需发车 237 次。

同理,我们可以求得下行方向的车次安排,见表 3。

表 3　下行方向的各个车站的转移客流量

站点	A0	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
1	22	29	48	**56**	47	39	31	-3	17	-23	-38	-101
2	795	869	1014	**1062**	975	878	609	515	509	299	212	-34
3	2328	2416	2695	**2763**	2460	2143	1586	1391	1365	1053	953	623
4	2706	2809	3123	**3193**	2878	2607	2154	1992	1986	1740	1685	1445
5	1556	1603	1770	**1812**	1654	1468	1183	1090	1087	925	888	732
6	902	945	1064	**1087**	1000	873	693	638	628	512	486	365
7	847	882	956	**985**	914	830	695	641	626	534	511	415

续表

站点	A0	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
8	706	726	805	**829**	761	688	576	543	544	455	435	339
9	770	794	880	**897**	828	752	628	592	603	501	476	373
10	839	889	998	**1018**	930	846	681	633	653	541	506	389
11	1110	1172	1303	**1322**	1225	1085	840	751	758	589	550	354
12	1837	1927	2181	**2223**	2045	1813	1378	1233	1246	964	903	566
13	3020	3163	3541	**3595**	3318	3025	2482	2309	2346	2030	1931	1612
14	1966	2084	2333	**2384**	2114	1849	1536	1415	1433	1243	1181	994
15	939	957	1060	**1089**	972	864	711	659	661	563	535	434
16	640	672	756	**780**	704	622	497	456	453	364	337	250
17	636	671	751	**763**	687	601	511	478	471	407	387	327
18	294	301	328	**332**	299	262	227	211	209	189	183	168

表 4　下行方向的各个时区的车次安排

时区	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
车次	6	9	24	27	16	10	9	7	8	9	12	19	31	20	10	7	7	6

可见,总共需发车 237 次。

公交车时刻表的安排:由于我们假设在各个时区内发车间隔相等,所以由以上的发车次数我们很容易得到始发站的发车时刻表。(见附录)

需要车辆数的求解:

知道了具体的发车时刻,我们就可以去求至少需要多少辆车。

设 D_1 为 A13 站必须的车辆数目,D_2 为 A0 站必须的车辆数目,$d_{in}(i)$为 A_{13}站在 i 时刻发车前由 A_0 已经开到的车次数,$d_{out}(i)$为 A_{13}站在 i 时刻发车前开出的车次数,$f_{in}(j)$为 A_0 站在 j 时刻发车前由 A_{13}已经开到的车次数,$f_{out}(j)$为 A_0 站在 j 时刻发车前开出的车次数。

目标函数: $\min \quad D_1 + D_2$

$$\text{s.t.}: D_1 + d_{in}(i) - d_{out}(i) \geqslant 1$$

$$D_2 + f_{in}(j) - f_{out}(j) \geqslant 1$$



（i,j 分别是已经确定的两个始发站的发车时刻。）

通过对这个线性规划求解，我们可以得出需要的最少车辆 $D=D_1+D_2=57$。

2. 模型二：模拟搜索

我们又用模拟的方法编了一个搜索程序去搜索可行的车次，我们尽可能地去模拟每步的情况，对各种情形做出分析。

用这种方法我们得出上行情况的一组车次安排（下行情况略）。

时区	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
车次	7	26	41	23	13	10	13	10	9	8	9	18	23	8	6	6	6	6

可见，总共需发车 242 班。

比较这两种方法得出的结果可以看到相差并不大，用模拟搜索总数要大一点，这是由于两种方法处理乘客等待时间时要求不一样造成的。

六、模型的评价及改进方向

1. 评价

我们对这种发车车次安排进行了评价，我们用平均等待时间和最长等待时间来评价乘客的满意度，用平均满载率来评价公交公司的利益。

平均等待时间 $\bar{t}$——所有乘客等待时间的平均值，

以上行方向为例，可以证明 $\bar{t}=\dfrac{\sum(t_i\times m_i)}{2\times\sum m_i}$，

其中：t_i 为第 i 时区发车间隔；m_i 为第 i 时区发出的车在全程运载的总人数。

平均满载率 $\bar{\beta}$ 为在这一天中，发出车的载客能力利用率。

同样，以上行方向为例，$\bar{\beta}=\dfrac{\sum m_i}{B\times100\times13}$

其中，B 为全天中上行方向总的发车班数。

用编程方法算得 $\bar{t}=2.17$ 分钟，$\bar{\beta}=81.58\%$。

(1)由于我们模型严格的时间控制,使得在整个运送过程中,没有乘客等待时间超过10分钟,并且平均等待时间 $\bar{t}$ 达到一个很小的值,对乘客来说,是非常满意的,这也有利于提高公司的社会效益。

(2) 但是对于公交公司来说,虽然车的载客能力利用率 $\bar{\beta}=81.58\%$ 已经较大,在早晨与晚上的时区内,我们安排的车的满载率远远低于50%,因此我们可以适当地减少早上、晚上的发车班数,以提高车的利用率。严格保证乘客的利益,是需要总的车辆数较大的主要原因。

2. 改进

(1)按照经验,在早高峰期间一般还存在着一个超高峰,我们可以增加高峰期间的发车班数;

(2)减少某些时区的发车次数;

(3)为了便于调度,我们还可以在不影响载客的前提下对一个小时内的发车次数进行微调,考虑到实际情况是一天的总客流量,向上的与向下的应该是相等的,为了不跑空车,我们还应保证一天中上行的总发车次数与下行的总发车次数相等;根据以上几点原则,我们可以进行手调,得到以下的发车方案(表5,表6)

表5 改进后的上行方向发车方案

时区	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
车次	6	24	40	24	15	10	10	10	8	8	8	15	24	8	5	3	3	2

表6 改进后的下行方向发车方案

时区	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
车次	5	8	24	30	15	8	8	6	8	8	10	20	30	20	8	6	6	3

每个始发站共需发车223次。

3. 进一步改进的方向

(1)每一个站点都可以看作是一个服务系统,可以近似看作是一个排队问题,整个问题就可以看作一个多服务系统的排队问题;



(2)找到一个更好的评价函数,能够同时对公司的利益与乘客的满意度进行比较;

(3)求发车次数时可以考虑的时间长一点,即把不同的时区合在一起,可以减少发车次数;由于时间紧张,我们没能继续深入研究这些问题。

七、模型的推广及应用

1.为更好地设计调度方案,采集数据使用的几种方法

(1)我们可以连续二或三天进行调查,在全天的运营调查中,调查各站所有乘客的下车地点和他们分别坐了几站;

(2)调查节假日的客流情况;

(3)各站下车乘客他们分别是从哪些车站上车的。

2.应用

实际的公交系统中是多条路线在运营的,每条路线都可以用这种方法来考虑,只是要多考虑一下各条路线之间的联系,因为它们有的路线是重合的。

八、参考文献

[1]同济大学.城市道路与交通.北京:中国建筑工业出版社,1979

[2]Eva part-enander MATLAB 5 手册.北京:机械工业出版社,2000

附录1　两个始发站的发车时刻表

附表1　A 13站点的发车时刻表

5:00:00	7:26:24	9:09:14	13:26:40	17:23:33
5:10:00	7:27:52	9:13:51	13:33:20	17:26:10
5:20:00	7:29:20	9:18:28	13:40:00	17:28:47
5:30:00	7:30:48	9:23:05	13:46:40	17:31:24
5:40:00	7:32:16	9:27:42	13:53:20	17:34:01
5:50:00	7:33:44	9:32:19	14:00:00	17:36:38
6:00:00	7:35:12	9:36:56	14:07:30	17:39:15
6:02:30	7:36:40	9:41:33	14:15:00	17:41:52
6:05:00	7:38:08	9:46:10	14:22:30	17:44:29

续表

6:07:30	7:39:36	9:50:47	14:30:00	17:47:06
6:10:00	7:41:04	9:55:24	14:37:30	17:49:43
6:12:30	7:42:32	10:00:00	14:45:00	17:52:20
6:15:00	7:44:00	10:05:27	14:52:30	17:54:57
6:17:30	7:45:28	10:10:54	15:00:00	17:57:34
6:20:00	7:46:56	10:16:21	15:07:30	18:00:00
6:22:30	7:48:24	10:21:48	15:15:00	18:07:30
6:25:00	7:49:52	10:27:15	15:22:30	18:15:00
6:27:30	7:51:20	10:32:42	15:30:00	18:22:30
6:30:00	7:52:48	10:38:09	15:37:30	18:30:00
6:32:30	7:54:16	10:43:36	15:45:00	18:37:30
6:35:00	7:55:44	10:49:03	15:52:30	18:45:00
6:37:30	7:57:12	10:54:30	16:00:00	18:52:30
6:40:00	7:58:40	11:00:00	16:03:20	19:00:00
6:42:30	8:00:00	11:05:00	16:06:40	19:10:00
6:45:00	8:02:37	11:10:00	16:10:00	19:20:00
6:47:30	8:05:14	11:15:00	16:13:20	19:30:00
6:50:00	8:07:51	11:20:00	16:16:40	19:40:00
6:52:30	8:10:28	11:25:00	16:20:00	19:50:00
6:55:00	8:13:05	11:30:00	16:23:20	20:00:00
6:57:30	8:15:42	11:35:00	16:26:40	20:10:00
7:00:00	8:18:19	11:40:00	16:30:00	20:20:00
7:01:28	8:20:56	11:45:00	16:33:20	20:30:00
7:02:56	8:23:33	11:50:00	16:36:40	20:40:00
7:04:24	8:26:10	11:55:00	16:40:00	20:50:00
7:05:52	8:28:47	12:00:00	16:43:20	21:00:00
7:07:20	8:31:24	12:06:00	16:46:40	21:10:00
7:08:48	8:34:01	12:12:00	16:50:00	21:20:00
7:10:16	8:36:38	12:18:00	16:53:20	21:30:00
7:11:44	8:39:15	12:24:00	16:56:40	21:40:00
7:13:12	8:41:52	12:30:00	17:00:00	21:50:00
7:14:40	8:44:29	12:36:00	17:02:37	22:00:00
7:16:08	8:47:06	12:42:00	17:05:14	22:10:00
7:17:36	8:49:43	12:48:00	17:07:51	22:20:00
7:19:04	8:52:20	12:54:00	17:10:28	22:30:00
7:20:32	8:54:57	13:00:00	17:13:05	22:40:00
7:22:00	8:57:34	13:06:40	17:15:42	22:50:00



续表

7:23:28	9:00:00	13:13:20	17:18:19	23:00:00
7:24:56	9:04:37	13:20:00	17:20:56	

附表 2　A0 站点的发车时刻表

5:00:00	8:19:57	11:26:40	16:22:10	18:15:00
5:10:00	8:22:10	11:33:20	16:25:20	18:18:00
5:20:00	8:24:23	11:40:00	16:28:30	18:21:00
5:30:00	8:26:36	11:46:40	16:31:40	18:24:00
5:40:00	8:28:49	11:53:20	16:34:50	18:27:00
5:50:00	8:31:02	12:00:00	16:38:00	18:30:00
6:00:00	8:33:15	12:08:35	16:41:10	18:33:00
6:06:40	8:35:28	12:17:10	16:44:20	18:36:00
6:13:20	8:37:41	12:25:45	16:47:30	18:39:00
6:20:00	8:39:54	12:34:20	16:50:40	18:42:00
6:26:40	8:42:07	12:42:55	16:53:50	18:45:00
6:33:20	8:44:20	12:51:30	16:57:00	18:48:00
6:40:00	8:46:33	13:00:00	17:00:00	18:51:00
6:46:40	8:48:46	13:07:30	17:01:56	18:54:00
6:53:20	8:50:59	13:15:00	17:03:52	18:57:00
7:00:00	8:53:12	13:22:30	17:05:48	19:00:00
7:02:30	8:55:25	13:30:00	17:07:44	19:06:00
7:05:00	8:57:38	13:37:30	17:09:40	19:12:00
7:07:30	9:00:00	13:45:00	17:11:36	19:18:00
7:10:00	9:03:45	13:52:30	17:13:32	19:24:00
7:12:30	9:07:30	14:00:00	17:15:28	19:30:00
7:15:00	9:11:15	14:06:40	17:17:24	19:36:00
7:17:30	9:15:00	14:13:20	17:19:20	19:42:00
7:20:00	9:18:45	14:20:00	17:21:16	19:48:00
7:22:30	9:22:30	14:26:40	17:23:12	19:54:00
7:25:00	9:26:15	14:33:20	17:25:08	20:00:00
7:27:30	9:30:00	14:40:00	17:27:04	20:08:34
7:30:00	9:33:45	14:46:40	17:29:00	20:17:08
7:32:30	9:37:30	14:53:20	17:30:56	20:25:42
7:35:00	9:41:15	15:00:00	17:32:52	20:34:16
7:37:30	9:45:00	15:05:00	17:34:48	20:42:50
7:40:00	9:48:45	15:10:00	17:36:44	20:51:24

续表

7:42:30	9:52:30	15:15:00	17:38:40	21:00:00
7:45:00	9:56:15	15:20:00	17:40:36	21:08:34
7:47:30	10:00:00	15:25:00	17:42:32	21:17:08
7:50:00	10:06:00	15:30:00	17:44:28	21:25:42
7:52:30	10:12:00	15:35:00	17:46:24	21:34:16
7:55:00	10:18:00	15:40:00	17:48:20	21:42:50
7:57:30	10:24:00	15:45:00	17:50:16	21:51:24
8:00:00	10:30:00	15:50:00	17:52:12	22:00:00
8:02:13	10:36:00	15:55:00	17:54:08	22:10:00
8:04:26	10:42:00	16:00:00	17:56:04	22:20:00
8:06:39	10:48:00	16:03:10	17:58:00	22:30:00
8:08:52	10:54:00	16:06:20	18:00:00	22:40:00
8:11:05	11:00:00	16:09:30	18:03:00	22:50:00
8:13:18	11:06:40	16:12:40	18:06:00	23:00:00
8:15:31	11:13:20	16:15:50	18:09:00	
8:17:44	11:20:00	16:19:00	18:12:00	

点评

本文针对公交车调度问题，首先定义两相邻正点之间的单位时间间隔为一个时区，并将问题适当地作了如下简化：将全天分为若干时区，每一个时区内发车时间间隔相等。这样使得模型的变量就可取为每一时区的发车时间间隔(发车总次数)，从而避免了以全天发车时刻为变量，大大降低了问题的维数、以及避免了变量个数不确定的情况。

文章认为一个时区内需要的车只与该时区所发车辆所转移的客流量有关。于是文章对每个站的客流量进行了处理，求出每个时段发的车的最大转移客流量，从已有数据集中得到一个新的数据集。

文章以满足乘客的等待时间要求为约束，以发车次数尽可能少，而且用到的车辆尽可能少为目标函数建立了一个单目标数学规划模型，求出了每个时区所发的车辆数。接着通过一个简单的线性规划模型，求出这条公交线路所需车辆最小值。然后综合考虑乘客的等待时间等一些因



素，给出各个时区的发车次数，进而确定全天发车时间表，由发车时间表求出需要的最少车辆数。

考虑到针对公交车调度问题是一个多目标问题，文章进一步提出了以下几个评价指标来讨论结果的合理性，以及对其他目标的满意程度：用平均等待时间和最长等待时间来评价乘客的满意度，用平均满载率来评价公交公司的利益。

最后，根据结果指出了进一步优化改进的方向。文章还用模拟的方法给出了一个方案，和数学规划模型的结果进行比较。

本文的成功之处在于：提出了时区的概念，然后通过时区重新划分客流，大大简化了问题的讨论；同时对于多个目标，文章以其中的一个为目标函数，一个为硬约束，其余为软约束(评价指标)。这两项处理使得建立的数学模型能够用来求解并得到结果。

此次竞赛中较突出的一个问题就是不少参赛队存在着建模与求解计算相脱节的现象，虽然建立了模型，但由于模型太复杂，无法用它来寻求答案，于是只好另起炉灶寻找答案，失去了建模的意义。将一个本来较复杂的模型，通过引入合理的假设加以简化，直到能够用计算工具处理，这是很重要的。当然这需要将得到的结果加以检验和分析，看看是否符合实际需要。

一般而言，对公交车的调度应以分钟为单位，而本文则以秒为单位，操作性稍差，应加以调整，这是本文的不足之处，也是其他许多论文存在的问题。

本文获得 2001 年全国大学生数学建模竞赛全国一等奖。

车灯线光源的优化设计

唐吉庆　史嘉凝　张皆喜

摘　要

本文首先建立了车灯光线反射的物理模型，通过严格的数学公式推导，尽可能用解析法来求得每一个物理量。在使用解析法困难的情况下，运用数值方法求得数值解，并且均在求解之后使用计算机数值模拟的方法来检验解的正确性。

对于问题一，本文建立了功率最小的非线性规划模型。通过线光源上的点和测试点的反射关系求得抛物面上反射点的解析解。利用光度学中“距离平方反比”模型求出点光源到测试点的光强度。最后，运用格点法搜索线光源长度，在满足光强要求的约束条件下求解规划模型，解得线光源长度 $I=3.88\text{mm}$。

对于问题二采用光迹追踪法，运用计算机模拟在抛物面上使用网格法遍历所有反射光线，并计算在测试屏上小区域的光通量分布，同时求得了测试屏上反射光的亮区和光强。

问题三在设计规范合理性讨论中，引入了一些重要指标，如纵向光线强度约束等，在量化指标后提出了方便实用的测量方法。

在求解过程中，我们大量使用解析法来推导公式，只在必要时用数值法代替。公式的推导和精炼使得在数值模拟时的搜索量大大降低，减少了计算机程序的求解时间。同时引入一些必要的近似方法来降低计算量，并在附录中给出了所使用近似方法的证明。



一、问题提出

(1)背景:在汽车工业中需要对汽车头部的车灯进行设计和测试。由于汽车前照灯的照明效果对夜间行车安全影响很大,因而对前照灯的光学性能提出了严格的标准。本文将对汽车头部车灯的光源进行优化设计和改进。

(2)重述:研究对象为灯罩形状旋转抛物面的汽车前照灯,车灯的对称轴水平指向正前方。在车灯的焦平面上,水平方向放置一条均匀分布的线光源。现有一套设计规范,为了测试车灯的反射光,在焦点正前方25米处的 A 点竖直地放置一测试屏,屏上引出一条水平的直线,沿直线方向在 A 点的同侧取 B 点和 C 点。符合设计要求:C 点的光照强度不小于某一额定值,B 点的光照强度不小于该额定值的2倍。

问题1 在满足该设计规范的条件下,计算线光源长度,使得线光源的功率最小。

问题2 针对得到的线光源长度,在有标尺的坐标系中画出测试屏上反射光的亮区。

问题3 讨论该设计规范的合理性。

二、符号约定

现建立三维直角坐标系如下(图1),以毫米为单位。

旋转抛物面 $\psi: y^2+z^2=2px$;

开口半径36,深度21.6,解得 $p=30$;

抛物面上的点 $P(x_p, y_p, z_p)$,$0\leqslant x_p\leqslant 21.6$,点 P 构成的集合:$\{P(l)\}$;

抛物面焦点:$F(p/2,0,0)$,焦距 $f=p/2=15$;

测试屏上点:$A(25015,0,0)$,$B(25015,1300,0)$,$C(25015,2600,0)$;

线光源:ξ,长度为 l $(0<l<60)$,中心为集点 $F(p/2,0,0)$;

线光源上的微元:$\mathrm{d}l$;

线光源上的微元看作点光源 $S(15, y_0, 0)$,简记为 $S(y_0)$;

抛物面焦点到光屏的距离 $D=25000$。

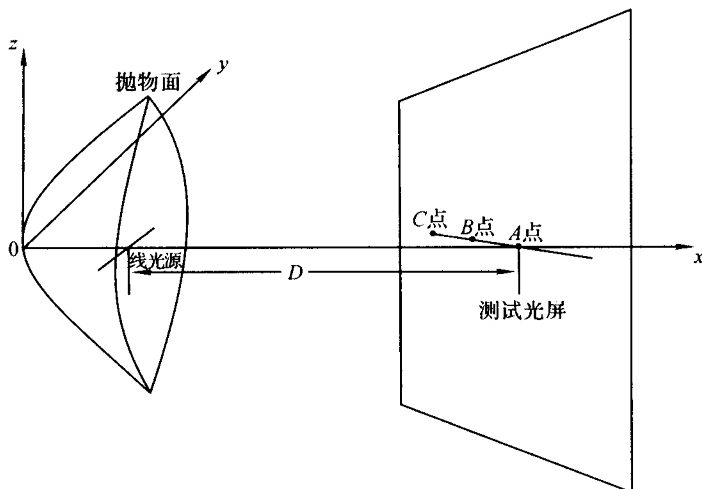


图 1　前照灯与测试光屏位置示意图

三、模型假设

由于该问题涉及到光学性能,我们需要对车灯的反射镜光学性质作必要的假设:

(1)反射镜内表面为严格的旋转抛物面;

(2)光线经过镜面发生全反射,没有折射和能量的损失。尽管工业设计中铝镀层反射系数大约为 90%,为了该问题的简化起见,认为反射系数为 100%;

(3)只考虑一次反射,不考虑二次或多次反射;

(4)线光源的光是均匀分布的,可以看成是点光源的叠加;

(5)由于抛物面的作用,使得反射光大大强于直射光,因此只考虑反射光在光屏上的强度。(证明见附录)

四、模型建立

问题一:规划问题

第一步　首先计算线光源上任意一个点光源在光屏上某定点(B 或



C)的光强度 $\mathrm{d}I(y_0)$。

第二步　对线光源 ξ 进行线积分，得到线光源在光屏上的某定点的光强度 $I(l)$。

第三步　提出规划模型，求解最小功率下的线光源长度。

设线光源长度为 L；

设满足 C 点的光强度不小于 1 的灯的最小功率为 $P_C(l)$；

设满足 B 的光强度不小于 2 的灯的最小功率为 $P_B(l)$；

$$\begin{cases} \min \max(P_C(l), P_B(l)) \\ \text{s.t.}: l \in (0,60) \end{cases}$$

问题二：根据问题一的线光源长度，用光线追踪法来绘制测试屏上反射光的亮区。

问题三：基于前两个问题结果的分析、评价和改进。

五、模型分析

(1)光度学基础

光强度的"距离平方反比"定理：

- 点光源射到某点的光强度与其距离 r 平方成反比，与点光源强度 I 成正比。
- 点光源到某一区域的光照度 A 还与其法线和入射光线的夹角 α 的余弦成正比。

公式描述为：

$$A = \frac{I}{r^2}\cos(\alpha)$$

由于本问题的特殊性，夹角的余弦可忽略不计。

(2)特点分析

根据该研究对象的特点，光学系统的精密性以及车灯设计安全性，用理论推导和软件计算相结合的方法，大量的计算都通过 MATLAB 和 C++ 语言进行推导和验证。在演算的过程中，每一个公式的推导都经过了理论和数值模拟的双重验证。我们在每一步都力求用解析的方法来求解，并通过数值模拟来检验；在无法求出解析值的情况下，再运用数值模



拟法。

六、模型求解

问题一：

(1)光线反射时入射向量、出射向量和法向量的关系

入射向量 $\vec{a}(x,y,z)$

出射向量 $\vec{b}(x',y',z')$

法向量 $\vec{N}(n_x,n_y,n_z)$

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1-2n_x^2 & -2n_xn_y & -2n_zn_x \\ -2n_xn_y & 1-2n_y^2 & -2n_yn_z \\ -2n_zn_x & -2n_yn_z & 1-2n_z^2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

(2)在线光源 ξ 上取点 $S(x)$，$-\frac{l}{2}<x<\frac{l}{2}$，取一小段微元 $\mathrm{d}l$，此时 $\mathrm{d}l$ 可以看作一个点光源，可计算得此点光源在光屏上任意点处的发光强度。

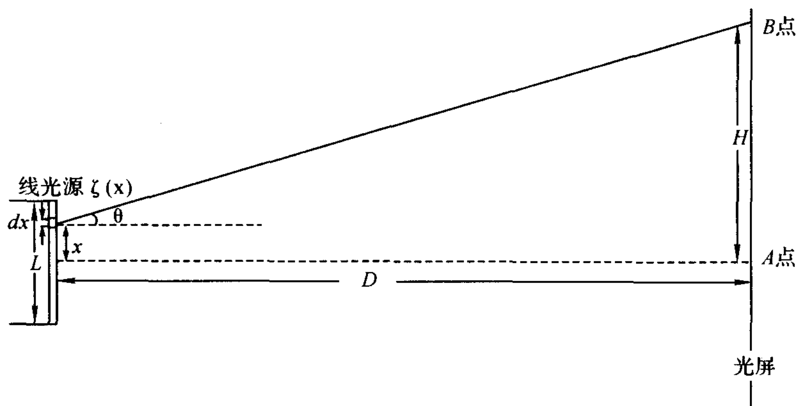


图 2　车灯与光屏的俯视图

①对给定的点光源和光屏上的接收点 B，求解抛物面的反射点。给定点光源 $P_0(15,P_{0y},0)$，光屏上接收点 $B(25015,1300,0)$，将抛物面上的反射点用参数方程表示为 $P_2(r^2/60,r\cos\theta,r\sin\theta)$。$r$ 为P_2 到 x 轴的



距离，θ 为 P_2 在 xy 平面内的转角，则入射向量为 $P_2 - P_0$，在 P_2 点的法向为 $n(30, -r\cos\theta, -r\sin\theta)$。

通过前述的旋转关系，我们可计算求出射向量 $\overrightarrow{b}$，则出射光线上的任何一点可以表示为 $P_2 + k \times \overrightarrow{b}$。

求解方程 $P_2 + k \times \overrightarrow{b} = B(25015, 1300, 0)$，可以得到以下两组解：

解法一：

当 $r\sin\theta \neq 0$，也即反射点不在水平面上，

$$k = -\frac{1}{2rP_{0y}\cos\theta}$$

$r = -10/(-2600 + p_{0y}) \times 3\hat{}(1/2) \times ((-2600 + p_{0y}) \times (7800 + 4997 \times p_{0y})) - (1/2)$

$\theta = a\cos(-125/13 \times p_{0y} \times 3\hat{}(1/2)/((-2600 + p_{0y}) \times (7800 + 4997 \times p_{0y})) - (1/2))$

反射点到光源的距离平方为：

$r_2 = 1/13 \times p_{0y} - 2 \times (8232380000 - 75100 \times p_{0y} + 13 \times p_{0y} - 2)/(-2600 + p_{0y}) - 2$

对光强 dI/r_2 积分值为：

$-338/31663/p_{0y} - 32539/401082276 \times \log(p_{0y})$
$+ 32539/8020364552 \times \log(8232380000 - 75100 \times p_{0y}$
$+ 13 \times p_{0y} - 2) + 40283217/271052230217116 \times 67591$
$-(1/2) \times$ ata
$n(1/168977500 \times (-75100 + 26 \times p_{0y}) \times 67591 - (1/2))$

对其满足 $0 \leqslant r_2 \leqslant 36$ 的 p_{0y} 值的范围（$-1.906 \sim -0.781$）

解法二：

当 $r\sin\theta = 0$，即反射点在水平面上时，

$k = -(r - 2600)/p_{0y}/(r - 2 - 900)$

$\theta = 0$

$p_{0y} = (r - 2600) \times (r - 2 + 900) - 2/(r - 4 - 1498200 \times r - 29360000 \times r + 1350810000)$

注:此时我们无法用 p_{0y}表示 r,因为出现了 r 的 5 次方程。但 p_{0y}可以表示成 $p_{0y}=f(r)$ 的形式,在积分时我们必须用 $p_{0y}=f'(r)\mathrm{d}r$,此积分值无法用解析式表示,故采用梯形公式数值积分进行计算。

反射点到光源的距离平方为:

$$r_2=(r-p_{0y})^2+(r\times r/60-15)-2$$

综合两种情况:我们可以根据 p_{0y}的值解出反射到 B 点的抛物面上的反射点,现将反射点的个数列式如图 3。

得到抛物面上的反射点 P 的分布,见图 3。

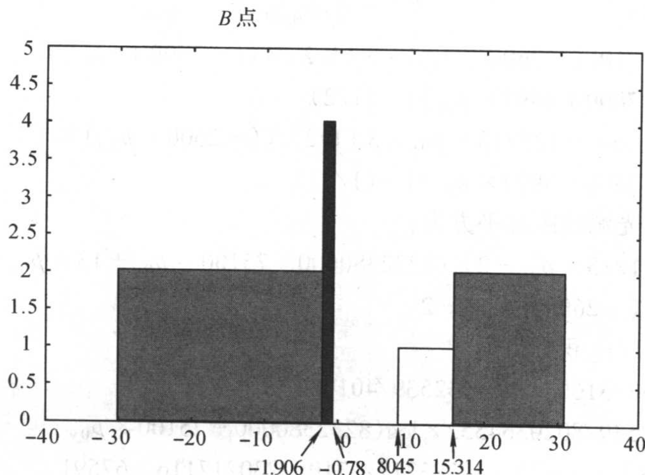


图 3 从线光源上 p_0 点出发经抛物面反射经过 B 点的反射点 p_2 数目分布

横坐标:p_0 点的 y 坐标 p_{0y}; 纵坐标:p 点的个数

用类似的方法可以解出反射到 C 点的情况,现将反射点的个数示于图 4。

②用“距离平方反比”求点光源在光屏上的光强度

$$\mathrm{d}I(l)=\frac{\mathrm{d}P}{4\pi r^2}=\frac{P\mathrm{d}l}{L}\times\frac{1}{4\pi|\overrightarrow{PS}|^2}$$

(3)沿线光源对光强度进行线积分

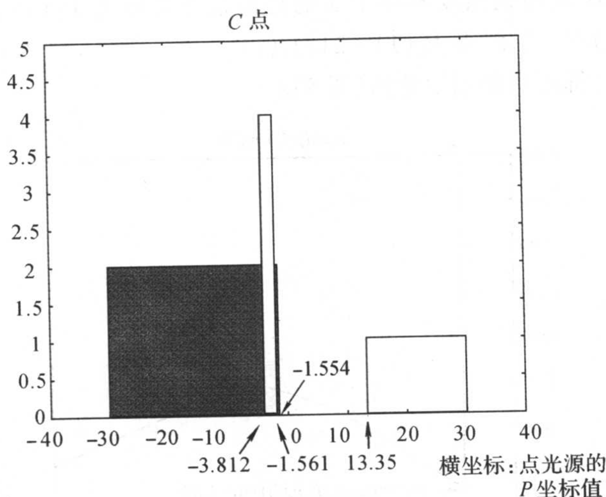


图 4　从线光源上 p_0 点出发经抛物面反射经过 C 点的反射点 p_2 数目分布

横坐标:p_0 点的 y 坐标 p_{0y}　　　纵坐标:反射点的个数

$$I(l) = \int_{-L/2}^{L/2} \mathrm{d}I(l) = \int_{-L/2}^{L/2} \frac{P}{4\pi L\,|\overrightarrow{PS}|^2}\mathrm{d}l$$

考虑反射点 $P(x_p, y_p, z_p)$:

①若 $z_p \neq 0$,则可以得到积分的解析值,

②若 $z_p = 0$,无法求得积分的解析值,所以用数值模拟来积分。

积分时的变量变换:

由于直接积分有困难,把积分变量改成 P 点的 y_p,不仅大大地简化了积分过程,而且可以在很短时间内算完。

(4)规划求解小功率下的线光源长度

我们把题中光强度的额定值取为单位 1,即要求 C 点的强度不小于 1,B 点的光强度不小于 2,求最小功率。

设　C 点积分所得值为 $J_C(1)$,B 点的积分所得值为 $J_b(1)$。

满足 C 点的光强度不小于 1 的灯的最小功率为 $P_C(1)$,则应有 $P_C(1) \times J_c(1)/1 \geqslant 1$,解得 $P_C(1) = 1/J_C(1)$。

满足 B 点的光强度不小于 2 的灯的最小功率为 $P_b(1)$，则应有 $P_b(1)\times J_b(1)/1\geqslant 2$，解得 $P_b(1)=21/J_b(1)$。

通过计算可以得到示意图(图 5)。

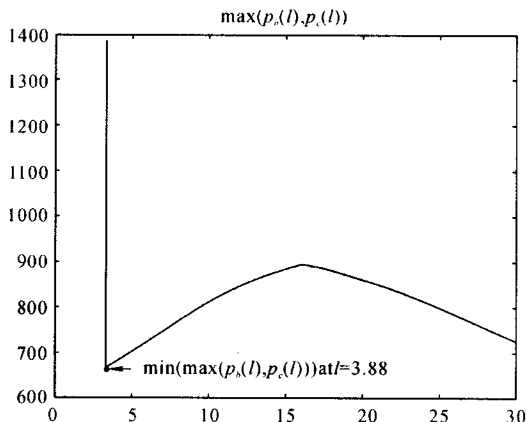


图 5

此问题可以转化为如下的解极小极大的非线性规划问题：

$$\begin{cases}\min\{\max(P_C(l),P_B(l))\}\\ \text{s.t.}\ P\in\Psi\ \text{即}\ l\in(0,60)\end{cases}$$

用格点搜索法解此规则，取线光源长度为 1，步长为 0.01，范围为(0,30)。可以搜出功率的变化规律，从而确定极小值为 $l=3.88$。

问题二：

利用光线追踪法来绘制测试光屏上的反射光亮区。

所谓光线追踪法，即运用计算机模拟追踪光线的走向，从而构造实际物体的三维照明模型。为此我们将测试屏划分为 10mm×10mm 的小区域，使用网格对整个抛物面进行光线投射模拟，对问题 1 所得到 3.88 的长度进行逐步模拟，即可求得测试屏上反射光的亮区和光的强度。

同时充分利用线性光源关于 y 轴和 z 轴的对称性，可节省 3/4 的运算量。

求得几个极值点如下：

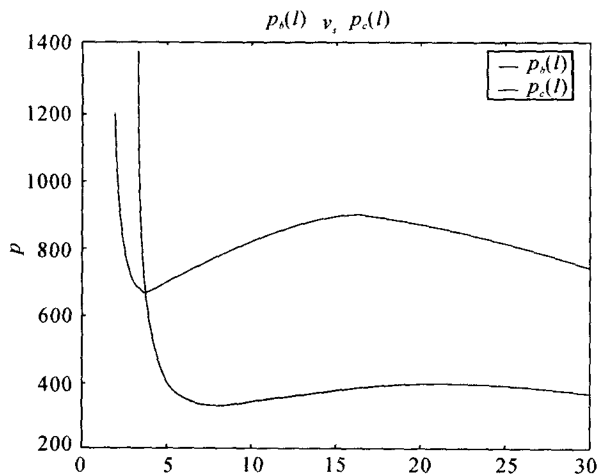


图 6

在 $x=25015$ 的平面内：

y 值最大为(25015,3252.14,0)；y 最小为(25015，-3252.14,0)。

z 值最大为(25015，±845.894,868.435)；z 值最小为(25015，±845.894，-868.435)。

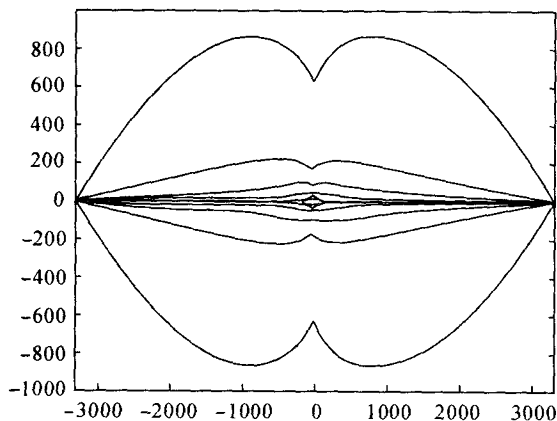


图 7 等高图

y 为 0 时，z 的取值范围为(－630.887,630.887)

将等高图示于图 7：

同时将所算光强度分布结果示于图 8，注意我们对光强度进行了对数变换，因为光强最大值和最小值相差太大。

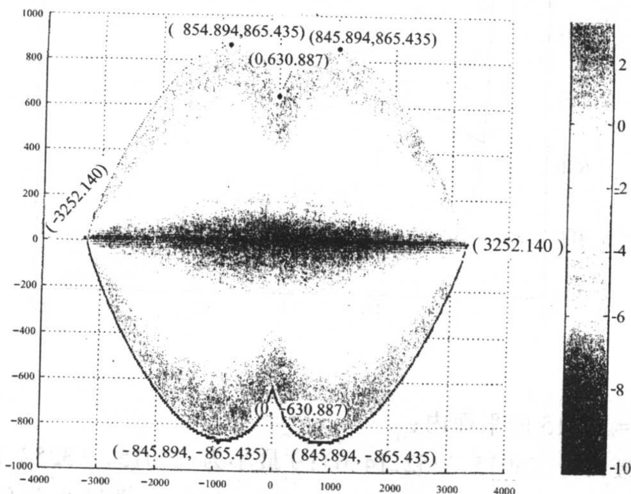


图 8(注：原图为彩色图)

从图上可以看到光在水平两侧对称分布，有唇形的等高线。

问题三：

我们查阅了前照灯的有关技术指标，参阅了中华人民共和国有关标准，前照灯的技术要求按 GB4785—1998 的规定，配光性能按照 GB4599—1997 的规定，类型和主要尺寸按照 GB/T4659—1984 的规定，使用和光束调整按照 GB7454—1987。

目前国际上有两套设计规范体系，美国的 SAE 体系和欧洲的 ECE 体系，美国的 SAE 体系习惯于测试远光灯，ECE 体系更习惯于近光配光。根据 GB4599—1997 标准，我国采用欧洲的 ECE 配光体系，前光灯配光试验的标准距离为 25 米。

(1)拟定设计规范的指标

①光分布的中间是高光强区域，以尽可能照亮前方的道路。



②光在水平方向的分布尽量分散，以尽可能照亮路边的情形。

③光尽量避免给对方司机造成眩光。

④光分布上：垂直方向上不要散得太开，要相对集中；

水平方向上要有足够的扩散；

水平方向还要有明暗截至线来防眩。

⑤两盏前照灯的总效果使得汽车正前方的灯光尽可能分布均匀。

⑥测量方法尽可能方便。

(2)各个指标的量化和分析

由于结果给出的水平方向的指标符合得相当好，这是线光源的缘故，也是题中所给设计规范合理的部分。

(3)各个指标的优化和改进

①为了在垂直方向满足光相对集中，防止眩光，我们在垂直方向再引入一个点 D，满足线光源在点 D 的光强度不大于某一额定值。

这样做的原因是相当明显的：为了保证水平方向的光尽量分散，规范中 B，C 两点，使那里的光强度不小于某些额定值；同理可推出，为了保证垂直方向的光尽量集中，可引入 D 点，使该点处的光强度不大于某额定值，防止对面司机的眩光，从而增强前照灯的安全性能。

②点 D 位置的确定：根据资料和估算，发现迎面而来的司机的眼睛高度比前照灯的线光源高度要高 0.375 米，因此初步确定 D 点的坐标为：$D(25015,1300,375)$。

D 点的光强度不大于某一额定值的$\sqrt{2}$倍，也即工程中的 3dB 点。同时，也应满足 B，C 两点原来的要求，保证水平方向的亮度。

③还可以进一步改进 B，C，D 点的位置，使得光强度更符合实际要求。

七、模型检验

我们以上的模型是通过严格的数学物理理论推导得到的，过程中涉及到许多方程的解析解与点的解析值。我们为了来证明建模的正确性，采用了两种途径来验证，一种是计算机的数值模拟，另一种是解析几何模型的建立。

1.非线性规划的计算机求解

费马定理:光线在实际路径上光程的变分等于零。

$$\delta(L) = \delta\int_P^Q n(x,y,z)\mathrm{d}s = 0$$

光线从一点 P 行进到另一点 Q 沿着光程 $\int_P^Q n(x,y,z)\mathrm{d}s$ 为极值的路径传播。因此,费马定理也称"光学极短时间原理"。

推论:光线从 S 点出发,经抛物面反射后,再到达光屏上 B 点的路线,必然是 S 点到抛物面上任意 P 点再到 B 点的光程中最短的道路。

证明:$\forall P$,S 点→P 点→B 点的光程 $\lambda(P)$ 为

$$\lambda(P) = \int_S^B n\,\mathrm{d}s = |\overrightarrow{SP}| + |\overrightarrow{PB}|$$

根据费马定理,

$\exists P_0, \text{s.t.}\ \lambda(P_0) = \min\lambda(P) = \min\{|\overrightarrow{SP}| + |\overrightarrow{PB}|\}$

从而我们得到了一种新的思路,可以用一个线性规划来进行抛物面上的反射点 P 的求解。

目标函数:$\min\{|\overrightarrow{SP}| + \overrightarrow{PB}|\}$

约束条件:P 点在抛物面 Ψ 上

具体数据代入,得到:

$$\begin{cases} \min\Big\{\sqrt{(x_P-15)^2+(y_P-l)^2+(z_P)^2} \\ \qquad +\sqrt{(x_P-25015)^2+(y_P-1300)^2+(z_p)^2}\Big\} \\ \text{s.t.}\ (x_P, y_P, z_P) \in \Psi: y^2+z^2=60x \end{cases}$$

我们在计算机上进行求解,可以自己编程对抛物面上的点进行搜索,也可以直接用 MATLAB 来求解。数值求解的结果和我们建模时理论推导公式得到的解析值吻合的很漂亮,从而验证了模型的正确值性。此外,我们也从另一个侧面得到了费马定理在抛物反射应用上的一个实例,并验证了其正确性。

2. 解析几何模型的检验

根据解析几何，我们知道，到两点距离之和等于某一定值的轨迹是一个椭球。那么把距离之和当作一个可变的参量，我们便可以得到一组以 S 点、B 点为焦点的椭球。当其中某个椭球与抛物面相切的时候，得到的切点便是我们所需要的 P 点集合。这个几何模型给我们的理论光学模型和非线性规划提供了一个形象的、可视化的数学依据，从而使得这个模型更有层次感。

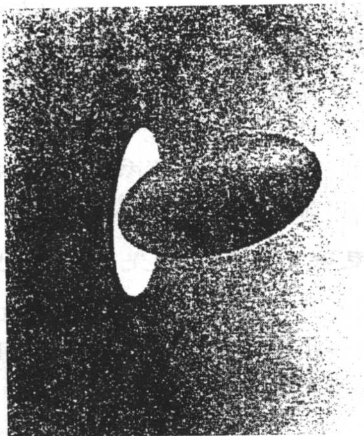


图 9

附录

1. 直射光线 $\{I_0\}$ 部分建模

数字模型公式 $\mathrm{d}I_0(l)=\dfrac{E}{[D^2+(H-l)^2]L}\mathrm{d}l$

然后沿着光源对上式进行线积分，

$$I_0(l) = \int_{L/2}^{L/2}\mathrm{d}I_0(l)$$

$$= -\frac{E}{DL}\arctan\frac{DL}{D^2+H^2-(L/2)^2}\int_{L/2}^{L/2}\frac{E}{[D^2+(H-l)^2]L}\mathrm{d}l$$

这样，我们就完成了 $\{I_0\}$ 部分未经抛物面反射的光线问题的求解。

本题中，证明了反射光线大大强于直射光线，因此本次建模中不考虑直射光线的影响。

2. "已知 S 点、B 点求 P 点"转化为"固定 P 点和 B 点反求 S 点"的意义

(1)问题提出是基于我们在运用前者时发现了直接求 P 点的解析解



是有一定困难的。求解 P 点坐标的方程高达5次，求得的解的个数也因 S 点的变化而变化。我们已经推算出在线光源 $\xi(l)$ 上，对应 P 点的解的个数随 l 呈分段分布，即在某一段线光源上，对应有抛物面 Ψ 上的 P 点个数不止一个。

(2)我们在以后的工作参见(6))会遇到关于线光光源的一个线积分。如果我们用前一种算法，那么我们在积分时将会遇到很大的困难。

3. 电磁学中关于光强度的定义

设点光源在某个方向上的立体角元 $\mathrm{d}\Omega$ 内发光通量元为 $\mathrm{d}\Phi$，则可定义电光源在此方向上的发光强度 I 为 $\mathrm{d}\Phi/\mathrm{d}\Omega$。根据经典电磁学，光强度是单位时间内单位面积上的能量。

$$E(x,t)=E_m\sin(kx-\omega t)$$

$$B(x,t)=B_m\sin(kx-\omega t)$$

$$I=\frac{\mathrm{d}U}{\mathrm{d}tA}=\frac{P}{A}=\frac{P}{4\pi r^2}=\frac{1}{\mu_0}E_{rms}B_{rms}$$

点　评

“车灯线光源的优化设计”来自汽车工业，所需光学知识较多，专业性强，要求参赛选手能在较短时间内熟悉并应用有关光学原理。数模竞赛选手来自各个专业，数模试题背景广泛，要求事先熟悉每一门学科是不现实的。因此可行的方法是培养在较短的时间内通过广泛查阅文献，阅读专业书籍，掌握关键知识，尽早进入问题情景的能力。

这篇论文首先建立了车灯光线反射模型。用微元法求得线光源上一点在光屏上任意点处的发光强度，再沿线光源对光强度进行线积分的办法求得给定点的反射光强度，利用非线性规划求得最优线光源长度。论文第二部分利用计算机模拟追踪光线的走向，绘制测试屏上反射光的亮区，最后根据汽车工业有关国家标准对该设计规范的合理性作了论证。论文结果叙述清楚，行文朴实流畅，提纲挈领，可读性强，在写法上也不乏可取之处。

对该题而言，尽管在建模整体思路上差别不大，不难得到与本文形式



相近的非线性规划,但两个光强度的具体表达式的推导和非线性规划的求解却有多种途径。对前者本文大量采用了解析法来推导公式。这样做有利于清晰揭示公式意义,缩短最终计算时间,特别是能减少数值计算带来的误差积累,能够实现形式计算功能的数学软件的使用为此提供了可能。对后者论文将遍历搜索和研究函数变化规律结合起来,从而大大减少了计算时间,准确地求得了全局最优解。文章在书写方面还存在一些缺陷,可以看出,作者在论文撰写时修改得还不够细致,这是本文的一个缺点。

本文获 2002 年全国大学生数学建模竞赛全国一等奖。

彩票中的数学

华　净　胡煜霄　徐　洁

摘　要

本文讨论了一个彩票种类设置的评价机制。我们推导出一种加权分析理论,通过权值对影响彩票发行的各个因子进行加权,得到相应的评价函数,用来评估抽奖方式和奖金分配方案对彩民吸引程度和管理部门的收益状况的影响。

我们先从概率统计的角度对评价的函数权值作出估算,继而用层次分析的方法导出权值,计算结果证明两种方法得到的对市面上常见的彩票方案的排序结果符合得相当好。最后得出所有现有方案中最优的方案。同时提出了搜索更好方案的算法。

合适的评价体系是本课题的关键。我们充分利用可供分析的数据,引入统计各方案在不同的权值组合下的排名情况,分析其排名的数学期望和方差等策略,最终得到各方案的优劣顺序。同时再一次验证了层次分析法所得到的权值的正确性。

在此基础上,我们对模型和彩票方案的前景作出了讨论与判断,并提出了无特别号码 33 选 7 的最佳方案。

最后我们用浅显生动的语言向公众介绍了我们的方案。

一、问题的重述

目前,市面上流行的彩票主要有“传统型”和“乐透型”两种类型。“传统型”采用“10 选 6 + 1”方案:先从 6 组 0～9 号球中摇出 6 个基本号码,



每组摇出一个,然后从 0～4 号球中摇出一个特别号码,构成中奖号码。投注者从 0～9 十个号码中任选 6 个基本号码(可重复),从 0～4 中选一个特别号码,构成一注,根据单注号码与中奖号码相符的个数多少及顺序确定中奖等级。“乐透型”有多种不同的形式,比如“33 选 7”的方案:先从 01～33 个号码球中一个一个地摇出 7 个基本号,再从剩余的 26 个号码球中摇出一个特别号码。投注者从 01～33 个号码中任选 7 个组成一注(不可重复),根据单注号码与中奖号码相符的个数多少确定相应的中奖等级,不考虑号码顺序。又如“36 选 6+1”的方案,先从 01～36 个号码球中一个一个地摇出 6 个基本号,再从剩下的 30 个号码球中摇出一个特别号码。从 01～36 个号码中任选 7 个组成一注(不可重复),根据单注号码与中奖号码相符的个数多少确定相应的中奖等级,不考虑号码顺序。

以上两种类型的总奖金比例一般为销售总额的 50%,投注者单注金额为 2 元,单注若已得到高级别的奖就不再兼得低级别的奖。其中一、二、三等奖为高项奖,后面的为低项奖。低项奖数额固定,高项奖按比例分配,但一等奖单注保底金额 60 万元,封顶金额 500 万元,各高项奖额的计算方法为:

[(当期销售总额×总奖金比例)－低项奖总额]×单项奖比例

本文根据这些方案的具体情况,综合分析各种奖项出现的可能性、奖项和奖金额的设置以及对彩民的吸引力等因素评价各方案的合理性,并且设计了一种“更好”的方案及相应的算法,并就此方案做出了具有实用意义和指导价值的前瞻性分析。

二、模型的条件和假设

(1)忽略模型的随机因素,以数学期望代替各参数的实际值。

(2)不考虑模型的时序特性。实际情况中,某期彩票的一等奖若未发出,其金额应累加到下一次开奖,我们在本模型中不考虑这一类特征。

(3)高级别奖项不再兼得低级别奖项。

(4)“传统型”彩票在购买时需要声明特别号码,而“乐透型”彩票无需声明特别号码。



三、符号的约定

P_1:一等奖金额；

λ_{p1}:一等奖开出率；

T_p:总中奖率；

$F(P_1,\lambda_{p1},T_p,P_{min})$:衡量对消费者吸引力的函数。

$P_{\min}$:最低奖的金额；

p_i:得 i 等奖的概率；

m_i:第 i 等奖的金额(当 i 为固定金额的奖项时)；

μ_i:第 i 等奖的比例(当 i 为按比例分配的奖项时)。

四、问题的分析与解答

1.常见彩票的概率分析

“乐透型”彩票规则分为两种,其一是 n 选 m,如题目例举的 33 选 7;其二是 n 选 $m+1$,如题目中所例举的 36 选 6+1。

对于 n 选 m 型选中 k 个,我们有

$$p_k=\frac{C_m^kC_{n-m}^{m-k}}{C_n^m} \tag{1}$$

若考虑是否选中特别号码,则有

$$p_k=\frac{C_m^kC_{n-m-1}^{m-k-1}}{C_n^m} \tag{2}$$

和

$$p_k=\frac{C_m^kC_{n-m-1}^{m-k}}{C_n^m} \tag{3}$$

对于 n 选 $m+1$ 型选中 k 个的情况,考虑是否选中特别号码,分别有

$$p_k=\frac{C_m^kC_{n-m-1}^{m-k}}{C_n^{m+1}} \tag{4}$$

和

$$p_k=\frac{C_m^kC_{n-m-1}^{m-k+1}}{C_n^{m+1}} \tag{5}$$

“传统型”彩票规则下,各奖项的中奖率的计算公式比较复杂,具体推

导在此从略。

根据如上所列公式,我们编写了 Matlab 程序,计算出了几种常见的彩票类型的中奖概率,如表 1 所示。

表 1　常见彩票类型的中奖概率

奖项 / 彩票序号	一等奖	二等奖	三等奖	四等奖	五等奖	六等奖	七等奖
1	2.00E-07	8.00E-07	1.80E-05	0.000261			
2	2.00E-07	8.00E-07	1.80E-05	0.000261	0.00342	0.042039	
3	2.00E-07	8.00E-07	1.80E-05	0.000261	0.00342	0.042039	
4	2.00E-07	8.00E-07	1.80E-05	0.000261	0.00342	0.042039	
5	6.41E-07	4.48E-06	9.42E-05	0.000283	0.002826	0.004709	
6	6.41E-07	1.41E-05	8.46E-0.5	0.000888	0.00222	0.0148	
7	4.91E-07	3.44E-06	7.56E-05	0.000227	0.002383	0.003971	0.026476
8	4.91E-07	3.44E-06	7.56E-05	0.000227	0.002383	0.003971	0.026476
9	4.91E-07	3.44E-06	7.56E-05	0.000227	0.002383	0.003971	0.026476
10	3.80E-07	2.66E-06	6.12E-05	0.000184	0.002021	0.003368	0.023572
11	3.80E-07	2.66E-06	6.12E-05	0.000184	0.002021	0.003368	
12	2.97E-07	2.08E-06	4.99E-05	0.00015	0.001722	0.00287	
13	2.97E-07	2.08E-06	4.99E-05	0.00015	0.001722	0.00287	
14	2.97E-07	2.08E-06	4.99E-05	0.00015	0.001722	0.00287	
15	2.34E-07	1.64E-06	4.10E-05	0.000123	0.001475	0.002458	
16	2.34E-07	1.64E-06	4.10E-05	0.000123	0.001475	0.002458	0.018843
17	1.86E-07	1.30E-06	3.38E-05	0.000101	0.001269	0.002115	
18	1.86E-07	1.30E-06	3.38E-05	0.000101	0.001269	0.002115	0.016916
19	1.49E-07	1.04E-06	2.81E-05	8.43E-05	0.001096	0.001827	
20	1.49E-07	1.04E-06	2.81E-05	8.43E-05	0.001096	0.001827	0.015224
21	1.49E-07	1.04E-06	2.81E-05	8.43E-05	0.001096	0.001827	0.015224
22	1.49E-07	1.04E-06	2.81E-05	8.43E-05	0.001096	0.001827	0.015224
23	1.49E-07	2.91E-05	0.001181	0.017051	0.10657		
24	1.20E-07	3.47E-06	2.08E-05	0.000292	0.00073	0.006566	0.008755
25	1.20E-07	3.47E-06	2.08E-05	0.000292	0.00073	0.006566	
26	1.20E-07	8.39E-07	2.35E-05	7.04E-05	0.000951	0.001585	0.013736
27	9.71E-08	6.80E-07	1.97E-05	5.92E-05	0.000828	0.00138	
28	2.61E-07	1.56E-06	5.16E-05	0.000129	0.002063	0.002751	
29	1.83E-07	9.16E-07	4.94E-05	9.89E-05	0.00262		

2. 奖项设置和奖金分配对彩民吸引力的影响

首先明确几个概念:

(1)一等奖的开出率:即一等奖的中奖概率,即表 1 中的第一列值。

(2)总中奖率:总中奖率指的是中奖总注数与总投注数的比例,即表1中每行的加和。

(3)一等奖的单注金额:由于彩民自己可决定所要的号码,每次抽奖所得的一等奖的个数是一个随机变量,由于所有中得一等奖的彩民平分一等奖金,那么一等奖的单注奖金也将随之变化。根据模型假设,用一等奖开出率乘以总人数为一等奖中奖人数的数学期望,再用一等奖总奖金额除以这个值,就得到了一等奖单注金额。

(4)最低奖金:指金额最小的奖金等级的奖金额。

(5)彩民种类的划分:传统经济学和消费心理学告诉我们"一般消费者在进行消费和投资时大都采取风险规避的态度,除非他们对未来的收益具有一定的预期"。然而我们看到,彩票作为一种特殊的商品,并不符合以上定律,甚至可以说是背道而驰。正如题中所述,无论最终的总投注数多少,彩票公司都要取其中的50%作为营销费用及利润,换句话说彩民在彩票的购买过程中必输。总的来说,每个彩民投资1元,只能收回0.5元。这不仅不符合"风险规避"的原则,甚至违反了"风险中性"的原则。可见用通常的概率论思想来考虑彩票购买中彩民的心理满意程度是不合理的。我们不妨将彩民分为"风险喜好","风险中性","风险规避"三类。"风险喜好"型的彩民会投入大量的资金去赌巨额奖金数的一等奖;而"风险规避"型的彩民则只是花几元到几十元钱去碰碰运气,"风险中性"则置于前两者之间。三类彩民对一等奖金额,一等奖开出率和总中奖率将有不同的喜好程度。此外这三类彩民在所有彩民中所占的比重也影响着彩票的发行状况。

3. 问题一模型的建立与评价

(1)对问题一的分析与建模

问题一要求我们评价各方案的合理性,解决问题的关键就在与建立合理的评价函数,分析出哪种方案对彩民的吸引力最大。由于影响彩民兴趣的因素很多,包括抽奖的方式以及各等奖的分配比例。我们从中归纳出三个最主要的因素:"一等奖单注金额";"一等奖抽出率";"总中奖率"。

按照正常人的心理特点,奖金额大的彩票一定更受欢迎,而奖金金额



大小的直观体现就是一等奖的奖金到底有多少。当然一等奖中奖概率越高也就越诱人，连续多次没有一等奖开出，彩民往往会兴趣下降。总中奖率和最低奖的金额则是保证该彩票能够吸引低端彩民的重要因素。因此我们认为 $F(P_1,\lambda_{p1},T_p,P_{\min})$ 是与 $P_2,\lambda_{p1},T_p,P_{\min}$ 正相关的，即

$$\frac{\partial F}{\partial P_1}>0;\quad \frac{\partial F}{\partial \lambda_{p1}}>0;\quad \frac{\partial F}{\partial T_p}>0;\quad \frac{\partial F}{\partial P_{\min}}>0$$

设 $F(P_1,\lambda_{p1},T_p,P_{\min})$ 的关系式为：

$$F(P_1,\lambda_{p1},T_p,P_{\min})=P_1^{\alpha}\lambda_{p1}^{\beta}T_p^{(1-\alpha-\beta)}P_{min}^{\gamma} \tag{6}$$

尽管是非线性关系，但只要对等式两边作对数运算即可得到：

$$\log F(P_1,\lambda_{p1},T_p,P_{\min})$$
$$=\alpha\log P_1+\beta\log\lambda_{p1}+(1-\alpha-\beta)\log T_p+\gamma P_{\min} \tag{7}$$

只要再以 $\log F,\log P_1,\log\lambda_{p1},\log T_p,\log P_{\min}$ 为新的自变量和因变量，就构成了一个线性方程。其中参数 $\alpha,\beta,1-\alpha-\beta$ 和 γ 分别表示“一等奖金额”、“一等奖抽出率”、“总中奖率”和“最低奖金”在彩民心目中的权重。

①模型的分析

下面我们对表 1 中的各个方案作出评价：

方法一：概率统计的方法

我们固定 α,β 和 γ 中的两个值，然后让另一个在一定的范围内变动，然后以一组 $(P_1,\lambda_{p1},T_p,P_{\min})$ 为样本，生成了一系列相对应的 $\log F(P_1,\lambda_{p1},T_p,P_{\min})$ 值。

产生 $(P_1,\lambda_{p1},T_p,P_{\min})$ 的方法是：先固定总的投注金额不变，如设定投注金额为 1000 万元（500 万注），除去 50% 划归彩票公司作为利润以外，剩余的 500 万元用于分发奖金。对每种抽奖方式来说，各奖项的概率是固定不变的，按照此概率计算出每一等奖项的中奖注数，以及一等奖项的单注金额。

计算公式为

$$\lambda_{p1}=p_1 \tag{8}$$

$$T_p=\sum p_i \tag{9}$$

$$P_i = \frac{(1 - \sum_{j=m}^{n} m_i p_j)\mu_i}{p_i} \quad (m, n \text{ 表示固定金额奖金的等级}) \tag{10}$$

按照上述计算公式,我们可以计算出表1中的29种不同彩票所对应的($P_1, \lambda_{p1}, T_p, P_{\min}$),参见表2。

表2 不同彩票所对应的判定参量

	一等奖单注金额	一等奖中奖率	总中奖率	最低奖金
1	2.47E+06	2.00E-07	0.00028	50
2	1.93E+06	2.00E-07	0.045739	5
3	2.09E+06	2.00E-07	0.045739	5
4	2.25E+06	2.00E-07	0.045739	5
5	7.56E+05	6.41E-07	0.0079166	5
6	6.59E+05	6.41E-07	0.018008	5
7	7.61E+05	4.91E-07	0.033137	5
8	9.45E+05	4.91E-07	0.033137	5
9	1.09E+06	4.91E-07	0.033137	5
10	7.95E+05	3.80E-07	0.0029208	10
11	1.70E+06	3.80E-07	0.0056359	5
12	1.77E+06	2.97E-07	0.004794	10
13	1.91E+06	2.97E-07	0.004794	10
14	2.05E+06	2.97E-07	0.004794	10
15	2.46E+06	2.34E-07	0.0040983	6
16	2.39E+06	2.34E-07	0.022942	5
17	3.14E+06	1.86E-07	0.00352	6
18	3.04E+06	1.86E-07	0.020436	2
19	4.29E+06	1.49E-07	0.0030366	5
20	3.38E+06	1.49E-07	0.018261	5
21	3.22E+06	1.49E-07	0.018261	5
22	4.39E+06	1.49E-07	0.018261	5
23	1.44E+06	1.49E-07	0.12483	2
24	4.21E+06	1.20E-07	0.016366	5
25	4.78E+06	1.20E-07	0.0076117	10
26	4.87E+06	1.20E-07	0.016366	5
27	5.47E+06	9.71E-08	0.002288	50
28	2.99E+06	2.61E-07	0.0049949	1
29	2.92E+06	1.83E-07	0.0027696	30

现在我们暂时不考虑最低奖奖金额的影响，即令 $\gamma=0$，而对 α,β 的不同组合值，得到一组 $\log F(P_1,\lambda_{p1},T_p,P_{\min})$ 的值。在这里 α,β 的值在一定程度上反映了彩民对风险的估计程度，如果某彩民的评价函数中的 α 值很大，那么他就属于"风险喜好"型，因为他对一等奖单注金额看得更重；反之则称"风险规避"型，这类人一般看重总的中奖率。我们以四组权值 $\alpha=0.6,\beta=0.1$；$\alpha=0.3,\beta=0.1$；$\alpha=0.2,\beta=0.4$；$\alpha=0.4,\beta=0.4$ 解出相应的 $\log F(P_1,\lambda_{p1},T_p,P_{\min})$ 值，并作出相应的直方图（图 1）。

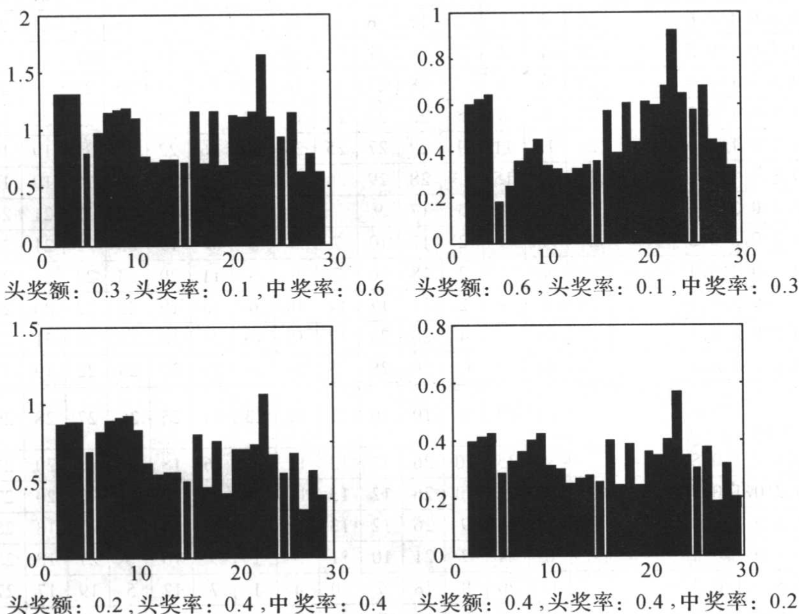


图 1　不同权值下各种彩票的评价函数值对比

从直方图中我们可以看出，第 23 种方案在这四种权值组合下都是最优的，但这一结果仅仅表示了以上的各方案在这四种权值组合下的优劣，而不能反映方案的总体情况，于是我们利用统计学的方法来弥补以上缺陷。

②模型的优化

先计算出表 1 中的各种彩票类型所对应的 $(P_1,\lambda_{p1},T_p,P_{\min})$ 值以及

$\log F(P_1, \lambda_{p1}, T_p, P_{\min})$在 $\alpha=0.1$ 到 0.8, $\beta=0.1$ 以及 $\alpha=0.1$ 到 0.6, $\beta=0.3$ 下的取值,然后将各类彩票在同一组权值下的结果相互比较并且排出名次:

表 3　各种彩票在不同权值下的排序表

$\alpha:\beta:(1-\alpha-\beta)$ \ 彩票序号	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0.1:0.1:0.8	29	4	3	2	17	14	7	6	5	8	19	23	22	21	24
0.2:0.1:0.7	29	4	3	2	18	16	7	6	5	9	19	23	22	21	24
0.3:0.1:0.6	29	4	3	2	19	16	10	6	5	15	20	23	22	21	24
0.4:0.1:0.5	29	4	3	2	26	17	14	13	12	15	19	25	22	20	23
0.5:0.1:0.4	29	6	3	2	28	21	15	14	13	16	24	26	25	23	22
0.6:0.1:0.3	29	10	6	5	28	27	20	17	14	23	24	26	25	22	19
0.7:0.1:0.2	28	13	11	9	29	27	25	24	20	26	22	23	21	19	18
0.8:0.1:0.1	23	17	15	13	28	29	26	25	24	27	21	22	20	19	18
0.1:0.3:0.6	29	6	4	3	17	9	7	5	2	8	18	23	22	21	24
0.2:0.3:0.5	29	4	3	2	17	10	7	6	5	8	19	23	22	21	24
0.3:0.3:0.4	29	4	3	2	18	16	7	6	5	11	20	23	22	21	24
0.4:0.3:0.3	29	4	3	2	21	17	14	10	6	15	19	25	22	20	23
0.5:0.3:0.2	29	10	6	4	28	25	17	15	13	24	18	26	23	21	22
0.6:0.3:0.1	26	16	14	13	29	28	25	24	21	27	20	23	22	19	17

$\alpha:\beta:(1-\alpha-\beta)$ \ 彩票序号	16	17	18	19	20	21	22	23	24	25	26	27	28	29
0.1:0.1:0.8	9	25	10	26	12	13	11	1	16	18	15	28	20	27
0.2:0.1:0.7	8	25	10	26	12	13	11	1	15	17	14	28	20	27
0.3:0.1:0.6	7	25	9	26	12	13	8	1	14	17	11	27	18	28
0.4:0.1:0.5	8	24	7	21	10	11	5	1	9	16	6	27	18	28
0.5:0.1:0.4	11	20	8	18	9	10	4	1	7	12	5	19	17	27
0.6:0.1:0.3	12	18	8	15	7	9	3	1	4	11	2	13	16	21
0.7:0.1:0.2	12	16	7	14	6	8	3	1	4	5	2	10	15	17
0.8:0.1:0.1	14	12	10	7	8	9	3	1	5	4	2	6	11	16
0.1:0.3:0.6	10	25	11	26	13	14	12	1	16	19	15	28	20	27
0.2:0.3:0.5	9	25	11	26	13	14	12	1	16	18	15	28	20	27
0.3:0.3:0.4	8	25	9	26	13	14	10	1	15	17	12	28	19	27
0.4:0.3:0.3	7	26	8	24	11	13	5	1	12	16	9	27	18	18
0.5:0.3:0.2	8	19	5	16	9	11	2	1	7	12	3	20	14	27
0.6:0.3:0.1	12	15	6	10	7	8	2	1	5	4	3	11	9	18



以上的 α,β 的取值范围是经过筛选的，去掉一些不合理的组合，保留一些常见组合的结果。对同一方案在不同权值下的名次进行统计，得出数学期望 E 和方差 σ：

E 可以用来衡量该方案的总体优劣程度，σ 则用来衡量方案的稳定性。

表 4 是各方案的 E,σ 值及其排名：

表 4　各种彩票的统计结果

彩票序号	名次均值	名次方差	均值排名	方差排名	彩票序号	名次均值	名次方差	均值排名	方差排名
1	28.29	1.73	29	5	16	9.64	2.21	8	8
2	7.57	4.77	5	18	17	21.43	4.67	23	17
3	5.71	4.34	3	15	18	8.50	1.83	7	6
4	4.50	4.09	2	14	19	20.07	6.70	19	23
5	23.07	5.25	26	20	20	10.14	2.48	9	10
6	19.43	6.72	18	24	21	11.43	2.28	12	9
7	14.36	7.29	15	26	22	6.50	3.96	4	13
8	12.64	7.46	13	28	23	1.00	0.00	1	1
9	10.71	7.09	11	25	24	10.36	4.94	10	19
10	16.57	7.44	16	27	25	13.29	5.43	14	22
11	20.14	1.96	20	7	26	8.14	5.40	6	21
12	23.86	1.41	27	4	27	21.43	8.16	23	29
13	22.29	1.33	25	3	28	16.79	3.47	17	12
14	20.64	1.15	21	2	29	24.64	4.50	28	16
15	21.86	2.66	24	11					

由上述统计结果可以看出，第 23 号彩票方案，即 35 选 7 无特别号的方案，无论在何种权值比例下都是最“好”的。这是因为它的一等奖奖金额和总体中奖率都很高，在投注总人数为 500 万人的情况下（投注总额 1000 万元，总奖金额 500 万元）分别为 428 万元和 12.5%，而只是一等奖的中奖率相对略低，为 1.487×10^{-7}，但平均每期能开出一等奖的概率仍有 74.36%。总体看来，35 选 7 的方案对彩民来说是有较大吸引力的。

对于 35 选 7 的方案，我们发现它的低奖项的奖金额很低，如五等奖的奖金仅有 2 元，恰为彩票的售价。这一点可能会对彩民的吸引力有负面的效应。为了消除这一效应，我们加入参数 $P_{\min}$，并将此方案的四、五等奖的奖金额分别改为 10 元和 5 元，即提高低奖项的奖金比例。通过对比修改前后的函数值，我们发现在调整之后，该方案的一等奖奖金额下降了 50% 左右，但是其综合吸引力并没有显著的下降，这也说明这种彩票形式是比较稳定的。

除了 35 选 7 无特别号的方案之外，其他的几种彩票方案的优劣程度可以由它们的名次均值来衡量。如传统型六等级方案也是比较好的方案。在排名均值比较接近的情况下，名次的方差反映了方案的稳定性，亦即不同倾向的人群在对头奖金额、头奖中奖率和总体中奖率的偏好程度有所不同时，方案能否比较稳定的吸引彩民。

方法二：层次分析法

以下我们用层次分析的方法来找出(7)式中的各权值。

目标层是“提高彩票对彩民的吸引力(O)”：准则层为“风险喜好型彩民的贡献($C1$)”，“风险中性彩民的贡献($C2$)”，“风险规避型彩民的贡献($C3$)”；措施层为“提高一等奖单注金额($P1$)”，“提高一等奖的开出率($P2$)”，“提高总中奖率($P3$)”，提高最低中奖金额($P4$)”。由此列出判断矩阵：

$$\begin{array}{cccc} O & C1 & C2 & C3 \\ C1 & 1 & 2 & 3 \\ C2 & 1/2 & 1 & 1/5 \\ C3 & 1/3 & 5 & 1 \end{array} \tag{11}$$

$$\begin{array}{ccccc} C1 & P1 & P2 & P3 & P4 \\ P1 & 1 & 2 & 3 & 4 \\ P2 & 1/2 & 1 & 3/2 & 2 \\ P3 & 1/3 & 2/3 & 1 & 4/3 \\ P4 & 1/4 & 1/2 & 3/4 & 1 \end{array} \tag{12}$$

$C2$	$P1$	$P2$	$P3$	$P4$
$P1$	1	1/3	2/3	4/3
$P2$	3	1	3/2	2
$P3$	3/2	2/3	1	4/3
$P4$	3/4	1/2	3/4	1

(13)

$C3$	$P1$	$P2$	$P3$	$P4$
$P1$	1	4/3	2/3	1/3
$P2$	3/4	1	1/2	1/4
$P3$	3/2	2	1	1/2
$P4$	3	4	2	1

(14)

完成层次单排序与一致性检验,并根据正互反矩阵及其特征根的性质,得到准则层各元素在目标层所占的权重和措施层各元素在准则层所占的权重。最后进行层次总排序和一致性检验,得到一组措施各元素对目标层的权重,即各措施"提高一等奖单注金额($P1$)","提高一等奖的开出率($P2$)","提高总中奖率($P3$)","提高最低中奖金额($P4$)"对增加对彩民吸引力的贡献:

$$\omega=(0.3345\quad 0.2836\quad 0.1964\quad 0.1855)$$

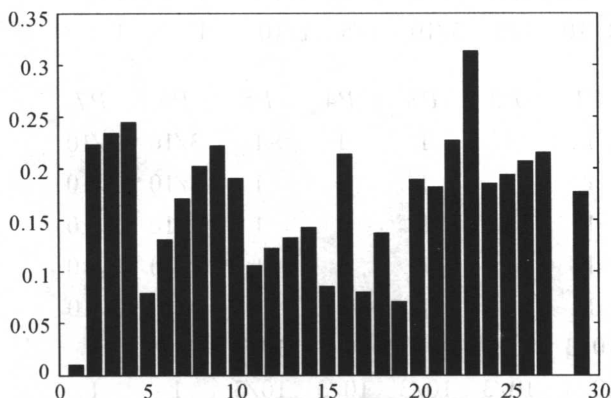


图 2　$\omega=(0.3345\quad 0.2836\quad 0.1964\quad 0.1855)$下的各方案对比

将上述这组权代入表达式(7),对每一组(P_1,λ_{p1},T_p,$P_{\min}$)得到相应的 F 函数,将这些函数值排序,图 2 是上述结果的直方图。

经过与表 4 所列的方法一中统计求数学期望相比较,结果基本相同。

下面进一步分析各等奖金分配与对彩民吸引力的影响。

在以上的层次分析法的基础上用措施层“一等奖的份额($P1$)”,“二等奖的份额($P2$)”,……,“七等奖的份额($P7$)”取代原有的措施层,得判断矩阵(这里仅列出了 ω_2 对应的判断矩阵):

O	$C1$	$C2$	$C3$
$C1$	1	5	1
$C2$	1/5	1	1/5
$C3$	1	5	1

(15)

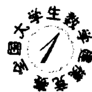
$C2$	$P1$	$P2$	$P3$	$P4$	$P5$	$P6$	$P7$
$P1$	1	2	3	2	1	10	10
$P2$	1/2	1	3/2	1	1/2	5	5
$P3$	1/3	2/3	1	2/3	1/3	10/3	10/3
$P4$	1/2	1	3/2	1	1/2	5	5
$P5$	1	2	3	2	1	10	10
$P6$	1/10	1/5	3/10	1/5	1/10	1	1
$P7$	1/10	1/5	3/10	1/5	1/10	1	1

(16)

$C3$	$P1$	$P2$	$P3$	$P4$	$P5$	$P6$	$P7$
$P1$	1	1	1	1	1	3/10	3/10
$P2$	1	1	1	1	1	3/10	3/10
$P3$	1	1	1	1	1	3/10	3/10
$P4$	1	1	1	1	1	3/10	3/10
$P5$	1	1	1	1	1	3/10	3/10
$P6$	10/3	10/3	10/3	10/3	10/3	1	1
$P7$	10/3	10/3	10/3	10/3	10/3	1	1

(17)

完成层次单排序与一致性检验,并根据正互反矩阵及其特征根的性



质，得到准则层各元素在目标层中所占的权重和措施层各元素在准则层中所占的权重。最后进行层次总排序和一致性检验（注：构造矩阵时可以利用传递闭包算法使得矩阵满足传递性），得到一组措施层各元素对目标层的权重，即各等级奖金分配对增加彩民吸引力的贡献：

当三种人群权重取 3:1:5 时，所得到的 ω_2：

$$\omega_2=(0.3830\quad 0.0819\quad 0.0753\quad 0.0819\quad 0.1018\quad 0.1380\quad 0.1380)$$

这样的彩民结构中“风险中性”的彩民数量很少，“风险喜好”型和“风险规避”型几乎各占一半，是一个较均衡的彩民结构。

在权重比为 10:1:2 的彩民结构下，得到

$$\omega_1=(0.5218\quad 0.0739\quad 0.0688\quad 0.0739\quad 0.0891\quad 0.0863\quad 0.0863)$$

显然这是“风险喜好”为主的彩民结构。

在权重比为 5:3:15 的结构下，得到

$$\omega_3=(0.1814\quad 0.1065\quad 0.0570\quad 0.0644\quad 0.0866\quad 0.1730\quad 0.3311)$$

这是“风险规避”者为主导的彩民结构。

方法二的结论：

权重比为 5:1:5，得到

$$\omega_2=(0.3830\quad 0.0819\quad 0.0753\quad 0.0819\quad 0.1018\quad 0.1380\quad 0.1380)$$

权重比为 10:1:2，得到

$$\omega_1=(0.5218\quad 0.0739\quad 0.0688\quad 0.0739\quad 0.0891\quad 0.0863\quad 0.0863)$$

权重比为 5:3:15，得到

$$\omega_3=(0.1814\quad 0.1065\quad 0.0570\quad 0.0644\quad 0.0866\quad 0.1730\quad 0.3311)$$

以上是针对不同的彩民结构，利用层次分析法解出的各奖项金额设置对彩民吸引力的权重，对于抽奖方式确定的情况，一个较好的奖金分配方案应尽量的接近以上权值。其总体趋势是“两头大，中间小”，这是符合实际情况的。从彩民的心理上看，不外乎两种心态：寻求刺激，搏一搏高额奖金的心理；玩玩而已，希望捞回成本的心理。一方面希望一等奖的金额尽量高，另一方面又希望总的中奖率比较高。这两者又由彩民结构中三类彩民的比重决定，反映在 ω 的值上就是“风险喜好”型彩民占主导的彩民结构中，一、二等奖的权重较大；反之，“风险规避”型彩民占主导的彩

民结构中，低等奖项的权重要设得大一些。

利用方法二判断同样的抽奖方式下不同奖金分配方案产生的影响。评判的依据是方案所实际产生的奖金分配比例与以上三种权重的接近程度，实际的奖金分配比例的计算方法是各奖项的单注金额×各奖项的概率。以下以方案6,7,8,9,16,18,20,21,22,24,26为例(均为7个奖项)，计算出各方案的奖金实际分配比例：

方案7	0.3740	0.0863	0.1151	0.1135	0.1191	0.0596	0.1324
方案8	0.4644	0.0663	0.1327	0.0454	0.1191	0.0397	0.1324
方案9	0.5333	0.0711	0.1067	0.0454	0.0715	0.0397	0.1324
方案10	0.3024	0.0756	0.1260	0.0918	0.1010	0.0673	0.2357
方案16	0.5595	0.0746	0.1119	0.0614	0.0737	0.0246	0.0942
方案18	0.5650	0.0997	0.1662	0.0507	0.0634	0.0211	0.0338
方案20	0.5021	0.0717	0.1435	0.0422	0.1096	0.0548	0.0761
方案21	0.4790	0.0639	0.0958	0.0843	0.1096	0.0913	0.0761
方案22	0.6525	0.0816	0.0816	0.0169	0.0548	0.0365	0.0761
方案24	0.5038	0.0672	0.1008	0.1459	0.0730	0.0657	0.0438
方案26	0.5829	0.0833	0.1665	0.0352	0.0475	0.0158	0.0687

由统计数据知，以上的这些方案的奖金分配比例与ω_3都相去甚远，也就是说这些方案在“风险规避”的人群中是没有吸引力的。这也验证了我们在文章开头做的分析，即彩民有别于一般消费者、投资者的差别是他们大都是“风险喜好”型的“机会主义者”。可以采取求各方案与ω_1的欧氏距离的方法来为判定各方案在相应的彩民结构ω_1中的优劣。

$$r_{ij} = \sqrt{\frac{1}{n}\sum(z_{jk} - \omega_{ik})^2} \quad (其中\ z_{jk}\ 表示第\ j\ 种方案的第\ k\ 个分量) \tag{18}$$

根据问题一的要求，我们比较以下相同的抽奖方案下的不同的奖金分配方式对彩民吸引力的影响。



取 7/30 抽奖方式为例,分析不同分配方案及其 r_{ij} 值:

	一等奖比例	二等奖比例	三等奖比例	四等奖比例	五等奖比例	六等奖比例	七等奖比例	r_{ij}值
方案 7	65%	15%	20%	500	50	15	5	0.1716
方案 8	70%	10%	20%	200	50	10	5	0.1160
方案 9	75%	10%	15%	200	30	10	5	0.0836

r_{ij}值越大,则奖金分配方案与层次分析法的结果越符合、越合理。从分配比例和 r_{ij} 值的变化趋势中我们可以得出以下定性结论:

同样比例的奖金,放在一等奖中要比放在二等奖和更低一些的奖中更具吸引力。

这一点是符合常识和一般人的理解的,对“风险喜好”型的彩民一等奖的吸引力要远胜于二、三等奖,一般不会有人单单为了获二等奖而购买大量的彩票。因此在同一抽奖方式下且固定低等奖项奖金额的基础上,适当的增加一等奖比例可以增加对彩民的吸引力。

4. 模型的推广与应用

问题二的解决

第二问题要求设计一个“更好”的算法。首先,我们的想法是利用数学规划的方法求出概率的最优组合值。

下面,我们给出目标函数和限制条件:

$$\min \quad -\log F(P_1, \lambda_{p1}, t_p, P_{\min})$$

即

$$\min \quad -(\alpha \log P_1 + \beta \log \lambda_{p1} + (1-\alpha-\beta)\log T_p)$$

其中

$$\lambda_{p1} = p_1,$$

$$T_p = \sum p_i,$$

$$P_i = \frac{(1 - \sum_{j=m}^{n} m_i p_j)\mu_i}{p_i} \quad (m, n \text{ 表示固定金额奖金的等级})。$$

约束条件为:

$$0 < p_1 < p_2 < p_3 < \cdots < p_7 < 1;$$

$$6 \times 10^5 < P_1 < 5 \times 10^6;$$

$$m_4 < \frac{\mu_3}{p_3} < \frac{\mu_2}{p_2} < \frac{\mu_1}{p_1}。$$

规划变量是 p_i，目标函数是评价函数的值最大化，限制条件共有三个：

(1)越是高额的奖项概率应该越低，并且概率值应在 0,1 之间取值。具体搜索时 p_i 的上界可定为 0.2 左右，因为几乎所有的商业彩票的总获奖概率都不会超过 20%，这样定义可以大大减少搜索的工作量。

(2)根据题目的要求，一等奖单注金额有上下限制；

(3)一等奖的单注金额＞二等奖的单注金额＞三等奖的单注金额＞四等奖的固定金额。

我们首先对规划式子作出如下简化 $T_p = \sum p_i$，已知 $p_1, p_2, p_3 \ll \sum_{i=4}^{7} p_i$，故 $\sum_{i=4}^{7} p_i \approx \sum p_i$，用前者取代后者。这样，目标函数 $\alpha \log P_1 + \beta \log \lambda_{p1} + (1-\alpha-\beta)\log T_p$ 就不含有 p_2, p_3 了，变量个数减少为 5 个。

我们编写了 Matlab 程序对该非线性规划进行求解，所用的方法是改进的穷举法。由于 p_i 的数量级有较大的差异，在搜索时，对每一个 p_i 使用了不同的步长。为了加快搜索速度，使用了逐步细化的算法。先在整个空间内以较大的步长寻找可行域，再在找到的可行域中进行精密搜索。

单纯使用规划的方法并不能直接求得一个好的设奖方案，但是通过搜索，我们发现，二、三等的奖金额趋向于四等奖的奖金额。这是由于我们在目标函数中并没有出现二、三等奖的 p_2、p_3 变量，在这种情况下，为了使得一等奖的奖金额最大，从而使得整个目标函数值取得极大，p_2, p_3 自然趋向于它们所能取得的最小值。由这一现象我们得到启发，即在七等级的设奖方案中，二、三等奖在其中的重要性是有限的。因此，在必要的情况下可以对它们进行简化。于是提出我们的最优设奖方案只设立五等奖项，其中后四等奖为定额奖，而将定额奖外剩余的奖金部分全部作为一等奖奖金。以上的这些特性在问题一的层次分析法中的相同抽奖方式下的不同奖金分配时也得到了反映。等量的奖金设在一等奖时带来的吸引力要比设在二等奖时大。另一方面，问题一中的最优解方案 23 也是一等奖 100% 的设奖方式。



通过求解非线性规划,我们定性分析了奖项设立规则:

在同一抽奖方案下的最优奖金分配方案是将扣除固定额奖金后的所有奖金设为一等奖。

根据上述结论,我们确定了我们要找的“更好”的方案应类似于23号方案,奖金分5等,一等奖比例100%设奖。基于这些特性,我们编写了第二个搜索算法,用于搜索比23号方案更好的方案。设这个更好的方案是一个 n 个数中选 m 个数(无特别号)的“乐透型”方案。我们仍用(P_1, λ_{p1}, T_p, $P_{\min}$)这四个参数来描述一个方案,用 $F(P_1, \lambda_{p1}, T_p, P_{\min})$ 来衡量方案的优劣。搜索采取以下策略:

选择代表三种彩民人群结构的权值组合:

$\omega_1=(0.3345\quad 0.2836\quad 0.1964\quad 0.1855)$,代表“风险喜好”,“风险中性”,“风险规避”人数比重为6:3:2的彩民结构;

$\omega_2=(0.2274\quad 0.3726\quad 0.2232\quad 0.1768)$,代表“风险喜好”,“风险中性”,“风险规避”人数比重为3:6:2的彩民结构;

$\omega_3=(0.2182\quad 0.2400\quad 0.2255\quad 0.3164)$,代表“风险喜好”,“风险中性”,“风险规避”人数比重为2:3:6的彩民结构。

我们在搜索最优值时有如下法则:当新得到的方案(n, m)对于 ω_1, ω_2, ω_3 加权后的结果都比23号方案好时才选取它为优于23号方案的解,同时方案的各权(P_1, λ_{p1}, T_p, $P_{\min}$)应满足以下边界条件:

$P_1<5\times10^6$;

$\lambda_{p1}>0.5$;

$P_{\min}>2$。

同时,再设定搜索范围为 $n=15\sim40$; $m=5\sim7$。

在这种法则下得到的优于23号方案,程序的输出结果为

● 从33里取7个,一等奖概率2.3408E-007,

各奖奖金[2693198.00 50.00 10.00 5.00 2.00]

● 从34里取7个,一等奖概率1.85887E-007,

各奖奖金[3472921.00 100.00 20.00 5.00 2.00]

● 从34里取7个,一等奖概率1.85887E-007,

各奖奖金[3556081.00 50.00 10.00 5.00 2.00]



- 从 35 里取 7 个，一等奖概率 1.48709E-007，
各奖奖金[3529965.00 2000.00 10.00 5.00 2.00]
- 从 35 里取 7 个，一等奖概率 1.48709E-007，
各奖奖金[4539610.00 100.00 20.00 5.00 2.00]
- 从 35 里取 7 个，一等奖概率 1.48709E-007，
各奖奖金[4628790.00 50.00 10.00 5.00 2.00]

通过以上输出，我们得到了 6 组优化方案。其中 33 选 7 的方式是评价函数值在该区域内的最优解，它在一等奖平均奖金额高于 200 万元的前提下，一等奖中奖概率达到了 2.34×10^{-7}，即当有 500 万人投注时，一等奖开出的期望达到了 1.17 人，基本上达到"每期都有大奖中出"的水平，使得该方案对彩民的吸引力保持在一个较高的水平。v 和 v_0 分别为该方案和 35 选 7 的原方案在三种权值下的评价函数值。

$v=[0.16186,-1.142,-0.28185]$

$v_0=[0.16031,-1.1844,-0.30015]$

$v-v_0=[0.00155,0.0424,0.0183]$

我们基于搜索结论，为管理部门提出以下 33 选 7 不含特别号的抽奖方案，其奖金的分配为：二等奖 50 元、三等奖 10 元、四等奖 5 元、五等奖 2 元，一等奖为所有剩余奖金。

5. 问题三的解决

(1)新型的 33/7 抽奖方案介绍

为了宣传我们的研究成果，推广新的彩票形式，我们给报社撰写了如下稿件，以期让读者了解我们的新方案，并对之产生兴趣。

给广大彩民的一封信

亲爱的彩民朋友们：

我们郑重推出最新式的彩票玩法"33 选 7 大赢家"。

通过长期细致的市场调研，我们精心设计了新的玩法。

1. 不设置普遍反映规律难找的特别号码，降低难度，更是较大程度减轻您不得不采用复式投注的经济负担和两难抉择。

2. 先前传统型排列式(6+1)玩法属排列方式。一要号码正确，二是位置正确，两大条件缺一不可——难度好大。



3.新玩法除一等奖外均采用固定奖金额,一方面扩大了中奖面,另一方面保证了一等奖的高数额,使广大彩民有了更广阔的选择余地。

我们推出的不含特别号码的“33选7大赢家”比传统型容易玩得多,比大多数的组合式“乐透”玩法更具吸引力,中奖面达到了13%以上,单注一等奖金额超过269万元,平均每次开奖一等奖的抽中人数达到1.17人。三项指标均超过市面上常见的彩票方案。

只要是对彩票有所心得之人,几次摸索下来,几乎保您不赔。

五、进一步研究

1.用时序因素对原模型进行修正

我们在文章开头的假设中提到“模型不考虑时序效应的影响”,这为我们建立和求解模型带来了极大的方便,同时也忽略了实际彩票发行过程中的一些重要特点。在实际一期彩票发售过程中,一等奖很可能空缺,这时按照彩票发售的规则,这一期的一等奖金额应自动的转到下一期,这使得下一期的一等奖单注金额增加,在这种情况下,$P_1(t)$将成为$\lambda_{p1}(t-\lambda)$,$\lambda_{p1}(t-2)\cdots$的函数。彩民的兴趣值不但不会因为连续的一等奖空缺而降低,反而会不断升高。此方案对彩民吸引力的函数F与λ_{p1}的关系大致如图3所示:

起初,λ_{p1}值很小表示一等奖多次空缺,这使得彩民对累计一等奖的期待增加,F值相应增加,且累计次数越多,F增加得越快,即$\frac{\partial F}{\partial \lambda_{p1}}>0$,$\frac{\partial^2 F}{\partial \lambda_{p1}^2}<0$。随着$\lambda_{p1}$的继续增长,函数出现一个拐点,在这之后$\frac{\partial F}{\partial \lambda_{p1}}>0$,$\frac{\partial^2 F}{\partial \lambda_{p1}^2}>0$。因为当一等奖抽出率上升到一定程度时,由于彩民人数以及资源的有限性,彩民们的兴趣度将趋向一个常量。

另外,除了λ_{p1}之外,P_1,T_p,$P_{\min}$都应有相类似的时序性质。F的表达式应改为:

$$F(P_1(t+n),\lambda_{p1}(t+n),T_p(t+n),P_{\min}(t+n),t)$$

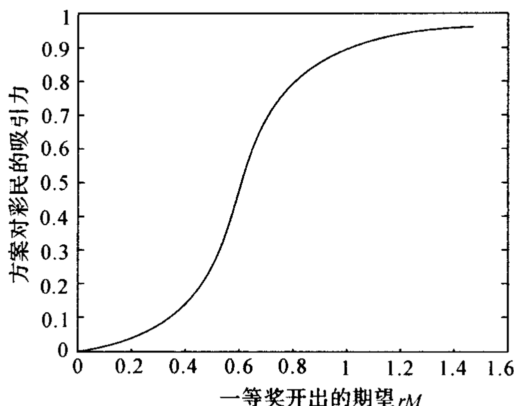


图 3

2. 问题一中的(18)定义是欧氏距离，可以用其他形式的距离代替

如：

$$\frac{\sum_{k=1}^{n}(x_{ik}-\bar{x}_i)(x_{ik}-\bar{x}_j)}{\sqrt{\sum_{k=1}^{n}(x_{ik}-\bar{x}_i)^2\sum_{k=1}^{n}(x_{ik}-\bar{x}_j)^2}}$$

还可以用指数相似系数法，算法平均最小法。

六、参考书目

[1]杨启帆，方道元．数学建模．杭州：浙江大学出版社

[2]贺仲雄．模糊数学及其应用．天津：天津科学技术出版社

[3]盛骤，谢式千，潘承毅．概率论与数理统计．北京：高等教育出版社

[4]王沫然．MATLAB 与科学计算．北京：电子工业出版社

[5]郎忠友．复合风险倾向函数构造和人们对风险态度的新解释．延边大学学报

[6]陈立基．彩票的商品特性及集市式销售的营销技巧．中国体育科技



点 评

彩票可以说是近年来风靡全国的一项新生事物。随着选号型彩票逐渐取代即开即得型成为彩票的主流，报刊和互联网上出现了一些无视彩票开奖中的随机性，以预测、选号为主要内容的所谓“数学模型”，一定程度上误导了广大彩民，而本题涉及的彩票方案设计问题却是数学模型可以真正发挥作用之处。

这篇论文首先利用古典概率对题中给出的彩票类型中获各项奖的概率作了正确的计算，尽管看似简单，却是进一步研究的基础。随后作者分别用概率统计的方法和层次分析法对各彩票方案作出评价，两种方法互相补充，互为印证。最后，作者尝试用非线性规划的方法寻求更好的彩票方案。论文作者敏锐地捕捉了彩票方案中四个影响彩民兴趣的因素：一等奖金额、一等奖中奖率、总中奖率、低奖金额，以及考虑风险喜好、风险中性、风险规避三种类型彩民对彩票方案的不同选择，这些都与命题者的初衷不谋而合。

与一些偏重于计算的赛题不同，本题的开放性较大，给选手以充分的发挥空间，但同时也容易导致泛泛而谈，浅尝辄止，研究难以深入。这就要求选手善于找出问题的关键因素，结合自身长处，选准突破口，呈现闪光点，到严格的推导和定量的计算贯穿建模过程，提高模型的说服力和可信度。

需要指出的是，对彩票方案的评价相当多的参赛队采用了层次分析法作为首选。诚然，经过数十年的发展和不断完善，层次分析法已成为将定性问题定量化的一种成熟的方法，但它也不可避免地带有许多主观的色彩。在竞赛短短 3 天的时间里，参赛队伍很难通过调查、咨询等方法得到权威的资料，因此需要慎重地使用。

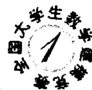
本文获 2003 年全国大学生数学建模竞赛全国一等奖。

露天矿生产的车辆安排

周　玥　陈　铭　卢军伟

摘　要

本问题研究的是露天矿生产中车辆的安排问题。题目要求我们根据不同的目标建立数学模型,给出一个班次的合理的生产计划及相应的石料运输安排,并用模型求解一个具体的生产实例。首先容易想到,这是一个数学规划问题。我们按照常规做法,列出问题的目标函数和约束条件,试图用现成的算法得到问题的解。但由于所要求解的车辆数是整数,而整数规划是相当难的,并且铲车数目的约束是非线性的,这样大规模的复杂计算很难用计算机实现,而且由于等待问题受许多随机因素的影响,这种"离散"规划模型很难处理好卡车等待时间的约束和随机问题,经过进一步的分析,考虑到一辆卡车往返运行周期远小于一个班次时间的实际情况,我们提出了一个基于网络运输流的"连续"模型,较好地解决了等待时间的问题。利用这个模型,我们化整数非线性约束规划为线性规划,得到了模型二的最优解:13 卡车,7 铲车,8.4829(万吨公里)总运量和模型三的最优解:20 卡车,7 铲车,10.3488 万吨总产量,4.9280 万吨岩石产量,14.6552(万吨公里)总运量。但在实际生产中,铲位和卸点一般不会在一个班次内连续不间断地工作。为了使模型的解更接近于实际生产情况,我们在上述模型求解过程中引入了各铲位(卸点)的工作饱和因子 P,对以上最优解进行了合理的调整,并通过图像分析,对 P 取不同值进行了灵敏度分析。考虑到应该使模型的应用具有尽可能的普遍性,我们还给出了一个车辆安排的快速算法,可用于非技术人员的一般性应用。此



算法速度很快，在奔三电脑上实现时间1秒不到。在进一步讨论的问题这一章节，我们分别对实时调度系统、铲车发生故障、多种车型、着眼于长期的规划问题等问题进行了讨论。

一、问题的重述

1. 背景

钢铁工业是国家工业的基础之一，铁矿是钢铁工业的主要原料基地。许多现代化铁矿是露天开采的，它的生产主要是由电动铲车（以下简称电铲）装车、电动轮自卸卡车（以下简称卡车）运输来完成。提高这些大型设备的利用率是增加露天矿经济效益的首要任务。

露天矿里有若干个爆破生成的石料堆，每堆称为一个铲位，每个铲位已预先根据铁含量将石料分成矿石和岩石。一般来说，平均铁含量不低于25%的为矿石，否则为岩石。每个铲位的矿石、岩石数量，以及矿石的平均铁含量（称为品位）都是已知的。每个铲位至多能安置一台电铲，电铲的平均装车时间为5分钟。

卸货地点（以下简称卸点）有卸矿石的矿石漏、2个铁路倒装场（以下简称倒装场）和卸岩石的岩石漏、岩场等，每个卸点都有各自的产量要求。从保护国家资源的角度及矿山的经济效益考虑，应该尽量把矿石按矿石卸点需要的铁含量（假设要求都为29.5%±1%，称为品位限制），搭配起来送到卸点，搭配的量在一个班次（8小时）内满足品位限制即可。从长远看，卸点可以移动，但一个班次内不变。卡车的平均卸车时间为3分钟。

所用卡车载重量为154吨，平均时速28公里/小时。卡车的耗油量很大，每个班次每台车消耗近1吨柴油。发动机点火时需要消耗相当多的电瓶能量，故一个班次中只在开始工作时点火一次。卡车在等待时所耗费的能量也是相当可观的，原则上在安排时不应发生卡车等待的情况。电铲和卸点都不能同时为两辆及两辆以上卡车服务。卡车每次都是满载运输。

每个铲位到每个卸点的道路都是专用的宽60米的双向车道，不会出现堵车现象，每段道路的里程都是已知的。

2. 问题

一个班次的生产计划应该包含以下内容:出动几台电铲,分别在哪些铲位上;出动几辆卡车,分别在哪些路线上,各运输多少次(因为随机因素影响,装卸时间与运输时间都不精确,所以排时计划无效,只求出各条路线上的卡车数及安排即可)。一个合格的计划要在卡车不等待条件下满足产量和质量(品位)要求,而一个好的计划还应该考虑下面两条原则之一:

①总运量(吨公里)最小,同时出动最少的卡车,从而使运输成本最小;

②利用现有车辆运输,获得最大的产量(岩石产量优先;在产量相同的情况下,取总运量最小的解)。

要求就两条原则分别建立数学模型,给出一个班次生产计划的快速算法。并针对一实例,给出具体的生产计划、相应的总运量及岩石和矿石产量。

二、模型的基本假设

(1)模型只考虑一个班次内的车辆安排。

(2)各卸点均须满足各自的产量要求和品位限制。

(3)各铲位的采装量必须小于其固定的最大石料产量。

(4)采矿过程中铲车、卡车均无发生故障,且卡车在各路线运行速度均相等,不会出现堵车现象。

(5)在一个班次内,卸点不可移动。

三、符号说明

Q'_{ij}: 在一个班次内从铲位 i 到卸点 j 运输的总矿石量;

Q''_{ij}: 在一个班次内从铲位 i 到卸点 j 运输的总岩石量;

L_{ij}: 铲位 i 到卸点 j 的距离;

Q_{ij}: 从铲位 i 到卸点 j 运输的总石料量;

P'_j: 卸点 j 对矿石量的产量要求;



P''_j：卸点 j 对岩石量的产量要求；

$B_{j\min}$：卸点 j 的品位限制的最小值；

$B_{j\max}$：卸点 j 的品位限制的最大值；

C_i：铲位 i 的矿石铁含量；

Q'_i：铲位 i 的矿石固定总数量；

Q''_i：铲位 i 的岩石固定总数量；

ΔT_{ij}：卡车在铲位 i 到卸点 j 之间的路线上循环运行一次需要的时间；

N_{ij}：铲位 i 到卸点 j 之间的路线上安排的卡车数量；

t_1：电铲的平均装车时间；

t_2：卡车的平均卸车时间；

V_{ij}：铲位 i 到卸点 j 之间卡车的车速。

四、问题的分析

该问题是一个多目标的整数规划问题。难点有好几处：

①如何解决卡车不等待的约束？

②初始铲位如何确定？基于原则一的模型，到底出动几台电铲？

③由于整数规划是相当难的，如何对大规模的整数规划求解？

我们先考虑建立问题的常规规划模型，试图用现成的算法得到问题的解。然后再考虑建立其他更适合的模型，并比较各模型的优劣。

五、模型的建立和求解

1. 基于原则一的模型

模型一（初始离散模型）

(1)模型一的建立

依据前面的分析，兼顾总运量与卡车的数量，我们建立如下的多目标规划模型：

目标函数：

① 总运量(吨公里)最少 $\min f=\sum_{i=1}^{n}\sum_{j=1}^{m}[(Q'_{ij}+Q''_{ij})\times L_{ij}]$

② 总车数最少 $\min\sum_{i=1}^{n}\sum_{j=1}^{m}N_{ij}$

约束条件:

① 对卸点的约束:

a.对产量的约束 $\sum_{i=1}^{n}Q_{ij}\geqslant P'_j+P''_j$;

b.品位限制 $B_{j\min}\leqslant\dfrac{\sum_{i=1}^{n}Q'_{ij}C_i}{\sum_{i=1}^{n}Q'_{ij}}\leqslant B_{j\max}$。

② 对铲位的约束:

a.矿石产量约束 $\sum_{j=1}^{m}Q'_{ij}\leqslant Q'_i$;

b.岩石产量约束 $\sum_{j=1}^{m}Q''_{ij}\leqslant Q''_i$。

③ 对卡车不等待条件的约束: $\dfrac{\Delta T_{ij}}{N_{ij}}\geqslant\max(t_1,t_2)$

④ 其他约束:

$$Q_{ij}=Q'_{ij}+Q''_{ij};$$

$$N_{ij},Q'_{ij},Q''_{ij}\geqslant 0;$$

$$\Delta T_{ij}=\frac{2L_{ij}}{V_{ij}}+t_1+t_2。$$

(2)目标函数说明

目标函数①使得在满足约束条件下卡车的总运量 f 最小。设有 n 个铲位 $i(i=1,2,\cdots,n)$,m 个卸点 $j(j=1,2,\cdots,m)$。

目标函数②使得在满足约束条件下,卡车的总量数最少。N_{ij} 表示铲位 i 到卸点 j 的路线上的卡车数量。

(3)约束条件说明

约束条件①中 a 式表明一个班次内各卸点必须满足其产量要求。b 式表明一个班次内各卸点必须满足其品位限制(岩石卸点的“品位限制”



即 铁含量可与矿石卸点的品位限制统一处理,详细说明将在模型的求解中给出)。$B_{j\min}$、$B_{j\max}$分别表示卸点 j 的品位限制的最小值和最大值。

条件②表明对各个铲位的采装均不能超过各铲位矿石、岩石的固定总数量。Q'_i、Q''_i 分别表示铲位 i 的矿石、岩石固定总数量。

条件③是在对每个铲位只分别考虑某一条路线上车辆的等待问题的假设下提出的。在任何一条从铲位 i 到卸点 j 的路线上,卡车的运行均不能发生等待现象。

(4)模型一的求解

由于每个铲位中既有矿石又有岩石,所以处理起来较为麻烦,因此我们提出将每个铲位按矿石和岩石分成两个不同的铲位,然后将它们按同等地位处理,这样就减少了一些复杂度。当然这样处理是要加条件限制的,其中我们运用了一个小技巧。岩石的铁含量是未知的,为了处理方便,我们可以假设岩石的铁含量是一个恒定的很大的负值,比如说 -10000%(这里 -10000%不代表现实意义,只是一个求解技巧),这样,岩石就不可能被送到矿石的卸点里,因为只要有一车这样的岩石,那么这个矿石卸点的"品位限制"肯定不能满足,如此例中的(29.5% ±1%)。同时我们又定义每个岩石卸点的"品位限制"即铁含量都为 -10000%,这样矿石就不可能被运到岩石卸点,因为它会使"品位限制"不为 -10000%。由此可见,上述定义可以放心地将岩石与矿石同等处理。

有了上述简化,我们先在 10 个铲位中选择 7 个铲位配铲车,这样有 $C_{10}^{7}=120$ 种情况,由于情况很少,我们对每种情况都求最优解,最后比较得出最终结果。我们不必再考虑 6 个或更少铲车的情况,因为它们其实都包含在 7 个铲车的情况中,相当于 7 个铲车中有的铲车没有用到的情况。

定了铲车后,这个问题就转化为一个单纯的多目标多约束的规划问题,这样就有了现成的算法和软件可以解决。但由于求解的时间复杂度较大,处理大规模的车辆安排问题时不是很快,所以我们又提出了以下的连续模型。

模型二(基于运输流思想的连续模型)

初始离散模型虽然直观易懂,但也存在着一些缺陷。按照初始模型

的思路，很难处理好各条路线上的卡车在一铲位汇合时铲位处的卡车等待问题。因为装卸时间和运输时间都不精确，随机因素影响比较大。另外，以上模型在计算机实现上也比较困难。

为了克服采用初始模型的缺陷，我们在此基础上引入了运输流的思想，建立了连续模型。

(1)模型二的建模准备

一般实际生产中，ΔT_{ij}与一个班次比起来均比较小。例如，实例中 $\max\{\Delta T_{ij}\} = \dfrac{2\max(L_{ij})}{V_{ij}} + t_1 + t_2 = 6.10 \times 2 \div 28 \times 60 + 5 + 3 = 34.14$(分)，而一个班次 T 为 8 小时。所以很容易将整个系统平均化处理，理解为一个网络运输流。若设卡车的满载量均为 q，则铲位 i 和卸点 j 之间路线上的一辆卡车的一次运行循环过程就可等价为一运输流，其流密度为$\dfrac{q}{\Delta T_{ij}}$。

(2)模型二的建立

目标函数：

$$①\min f = \sum_{i=1}^{n}\sum_{j=1}^{m}\left(\frac{N_{ij} \times q \times T}{\Delta T_{ij}} \times L_{ij}\right)$$

$$②\min \sum_{i=1}^{n}\sum_{j=1}^{m} N_{ij}$$

约束条件：

① 对卸点的约束：

a. 对产量的约束 $\sum_{i=1}^{n} Q_{ij} \geqslant P'_j + P''_j$

b. 品位限制 $B_{j\min} \leqslant \dfrac{\sum_{i=1}^{n} Q'_{ij}C_i}{\sum_{i=1}^{n} Q'_{ij}} \leqslant B_{j\max}$

② 对铲位的约束：

a. 矿石产量约束 $\sum_{j=1}^{m} Q'_{ij} \leqslant Q'_i$



b. 岩石产量约束 $\sum_{j=1}^{m} Q''_{ij} \leqslant Q''_{i}$

c. 对铲车的约束 $\sum_{i=1}^{n} F_i \leqslant n$ 且 $F_i = \begin{cases} 0 & \sum_j N_{ij} = 0 \\ 1 & \sum_j N_{ij} > 0 \end{cases}$

③ 对卡车不等待条件的约束：

a. 铲车处 $\sum_{j=1}^{m} N_{ij} \dfrac{q}{\Delta T_{ij}} \leqslant \dfrac{q}{t1}$

b. 卸点 $\sum_{i=1}^{n} N_{ij} \dfrac{q}{\Delta T_{ij}} \leqslant \dfrac{q}{t2}$

(3)目标函数说明

目标函数①即是在满足约束条件下，卡车的总运量最小。事实上，初始模型的 Q_{ij} 就等于 $\dfrac{N_{ij} \times q \times T}{\Delta T_{ij}}$，其中 T 为一个班次时间。

目标函数②使得在满足约束条件下，卡车的总量数最少。

(4)约束条件说明

条件①与初始模型处理相同。

条件②中增加了对铲位位置确定的约束，便于计算机实现时搜索判断实际使用铲位的数量和位置。当 $F_i = 1$ 时表示铲位 i 有电铲采装，当 $F_i = 0$ 时表示铲位 i 没有电铲采装。

条件③是与初始模型区别最大，也是此模型最具特色的地方之一。连续流量的约束巧妙地处理了卡车等待问题，避免了离散模型中需要考虑的随机性因素影响，使问题得到了很大的简化。

(5)模型二的求解

该模型是一个非线性约束的整数规划模型，考虑到整数变量有几十个，如果用整数规划来求解，计算量会十分庞大，计算机难以在短时间内实现。所以我们考虑用线性规划方法，得到松弛的最优解，再进行局部调整。

①多目标问题转化为单目标问题

我们将卡车数量的优化 $\min \sum_{i=1}^{n} \sum_{j=1}^{m} N_{ij}$ 作为一约束条件，在局部调整过

程中作为判断调度方案优劣的标准。这样,多目标问题就转化成了单目标问题。

②算法介绍

第一步:以不超过 n 个铲位为约束,各条线路上的卡车数量为求解量,卡车总运量为目标函数,应用线性规划解法求解[1]。

第二步:判断所用铲位数是否满足不超过 k 个.若满足,则得到了松弛的最优解。若不满足,则回到第一步。最坏情况下有 C_n^k 次计算量。(k 为铲车总数)

③模型结果分析

利用上述算法,易得到松弛的最优解(单位:辆)

	铲位1	铲位2	铲位3	铲位4	铲位8	铲位9	铲位10
矿石漏		0.8183			1.8199		0.3031
倒装场Ⅰ		1.0489		1.1583			
倒装场Ⅱ		0.7194					1.4865
岩石漏	1.8166		1.1821				
岩场						1.8183	0.3227
总运量	8.4829(万吨公里)			卡车数	12.4941		

对于车辆数的小数部分,一种简单易行的方法是直接进行手工局部调优。例如本例中,将各路线上的车辆数互相搭配,以相加趋近整数为目的,即可使得某些卡车可以去多个铲位与卸点,从而让资源得到充分利用。因为卡车切换到不同铲位时必须是从卸点过去的,一定是空车,所以这些可能多走的路程并不会给总运量带来影响。例如我们搭配如下(1.8166,1.1821),(2.2059,1.7585),(1.4865,0.4140),(0.8183,0.7194,0.7185,1.3211)。搭配后只需给 3+4+2+4=13 辆卡车即可。当然每个搭配中总和越趋近于整数,元素越少就越优,因为毕竟在满足卡车总数少时还希望卡车跑的铲区也尽可能少。

此外,我们根据这个目标提出了一个代替手工调优的机选算法:

Ⅰ.分别取各个路线上卡车辆数的小数部分,记为:$a_1,a_2,\cdots,a_n$

Ⅱ.若 a_i,a_j 满足 a_i+a_j 的小数部分在[0.5,1]区间内,则我们称



(a_i, a_j)为可行集 E 的一个元素。

Ⅲ.用深度优先法在 E 中搜索可以覆盖 $a_1, a_2, \cdots\cdots, a_n$ 中尽可能多元素且没有重复的子集,然后按子集中的元素相应的配对即可。

显然这里求的不是最优配对,但至少是比较优的,因为这对整个模型影响不是特别大,所以不必花很多精力,只从简单考虑。

然而配对这一方法也有不大完善的地方,分析如下:

● 若铲位(或卸点)的工作状态已达到饱和,即工作效率为100%,则与该处搭配的其他处卡车在此时来到便会发生等待现象。

虽然这种可能性不是很大,但是还是不能忽略的。

● 没有考虑到空车在改变运行线路时的损耗,可能会带来一定不合理性。

(4)模型的改进

为了解决上述解法产生的不合理因素,我们引入饱和因子 P 的概念。例如 $P=0.9$ 则表示规定铲车的计划时间利用效率不能超过0.9,剩下的0.1可以用于系统稳定性的考虑,如解决上述模型中不同的搭配可能所造成的等待问题,甚至还可以推广用于解决车的故障所带来的阻塞问题。而且在后面将看出饱和因子的小范围变化并不会过多地影响最优解,如 $P=0.9$ 与 $P=1$ 时所得的最优结果差别不大,但此时我们就有更多的备用时间用于解决随机出现的堵车问题。

我们可以设定每个铲位(卸点)的工作饱和因子 $P_i(P_j)$,于是卡车不等待的约束转化为:

① $\sum_j N_{ij} \frac{q}{\Delta T_{ij}} \leqslant P_i \frac{q}{t1}$

② $\sum_i N_{ij} \frac{q}{\Delta T_{ij}} \leqslant P_j \frac{q}{t2}$

在改进算法给出以前,我们先对饱和因子 P(将 P_i 和 P_j 等同对待)和目标函数 f 的关系进行分析,用 MATLAB 画图软件[2]作出两者的关系曲线,如图1所示:

从图像中我们可以看出总运输量 f 是饱和因子 P 的递减函数。同时当 P 小于0.78时,求不出问题的解。我们权衡了各方面要求,此例中取定 $P=0.9$,得到解如下:

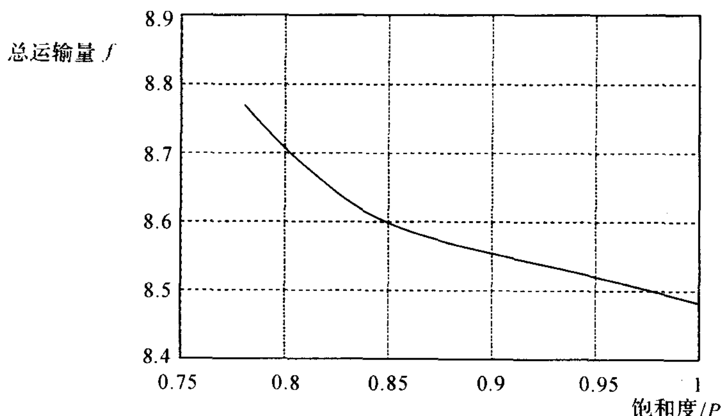


图 1

	铲位 1	铲位 2	铲位 3	铲位 4	铲位 8	铲位 9	铲位 10
矿石漏		0.8183			1.8199		0.3031
倒装场Ⅰ		1.0489		1.1583			
倒装场Ⅱ		0.7194					1.4865
岩石漏	1.8166	–	1.1821				
岩场						2.0692	0.1138
总运量	8.5554(万吨公里)			卡车数	12.5361		

将车辆的小数部分调优为整数,我们得到以下解(单位:辆次):

	铲位 1	铲位 2	铲位 3	铲位 4	铲位 8	铲位 9	铲位 10
矿石漏		13			55		10
倒装场Ⅰ		41		44			
倒装场Ⅱ		14					71
岩石漏	81		43				
岩场						79	6
总运量	8.595(万吨公里)			卡车数	13		

当然,一个好的生产计划还应该给出具体卡车的运行方案一个运行方案,如下:(数字为卡车的编号,括号中的数为该车的运行次数,如 6

(10)表示第 6 号车在此路线上运输 10 次)

	铲位 1	铲位 2	铲位 3	铲位 4	铲位 8	铲位 9	铲位 10
矿石漏		7(13)			8(29) 4(26)		12(10)
倒装场Ⅰ		3(39) 4(2)		5(37) 6(7)			
倒装场Ⅱ		6(14)					13(47) 12(24)
岩石漏	1(44) 2(37)		12(8) 11(35)				
岩场						9(38) 10(38) 7(3)	2(6)
总运量	8.595(万吨公里)			卡车数	13		

2. 基于原则二的模型

模型三

(1)模型三的建立

该模型是基于原则二建立的,其约束条件与模型二相同,只是目标函数变了。其目标函数为:

① $\max\sum_{i=1}^{n}\sum_{j=1}^{m}Q_{ij}$(总产量最大)

② $\max\sum_{i=1}^{n}\sum_{j=1}^{m}Q''_{ij}$(岩石产量优先)

③ $\min\sum_{i=1}^{n}\sum_{j=1}^{m}[(Q'_{ij}+Q''_{ij})\times L_{ij}]$(总运量最小)

(2)模型三的求解

①多目标问题的转化

此模型仍是一个多目标规划问题,我们采取约束法[3],以总产量最大作为主要目标,以岩石产量为自变量,在确定岩石产量下界情况下(如本例中岩石产量下界为 1.9 + 1.3 = 3.2(万吨)),用步长搜索法,通过描点法用 Matlab 作出了岩石产量与最优总产量关系的曲线,如图 2 所示:

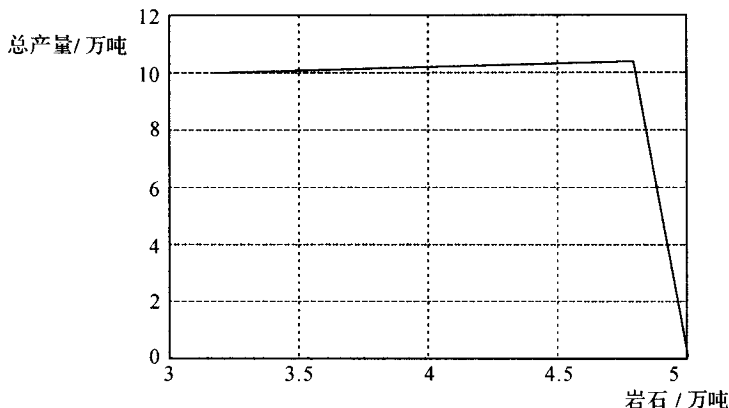


图 2

由上图可明显地看出，当岩石产量在 4.9280 万吨之前时，曲线非常平缓，可认为此时最优总产量变化不大，而当岩石产量达到 4.9280 时，曲线急剧下降，这是因为我们预先规定当任一约束条件不满足时，最优总产量的值均为零。故这个转折点即为我们所求的工作状态点，即保证最优总产量变化不大时，让岩石产量尽可能地大。

②算法介绍

当工作状态点确定后，该模型可以很好地沿用模型二的求解方法，求解时只需加入已知的固定岩石产量和最优总产量的约束条件，以总运量最小为目标函数，可求出模型的松弛最优解。

	铲位 1	铲位 2	铲位 3	铲位 4	铲位 8	铲位 9	铲位 10
矿石漏			1.6177		1.2702	0.2313	
倒装场Ⅰ	0.5730	1.7390	0.2219	1.8243			
倒装场Ⅱ			1.4226		1.4521		0.8222
岩石漏	1.7673	0.9008	0.8700	0.7312			
岩场				0.3250	0.4475	2.2385	1.2420
总产量(万吨/公里)	10.3488	总岩石产量	4.9280	总运输量	14.6552	车数	19.6967

③模型的改进

猜想:最优总产量10.3488万吨是在每个铲位均不间断工作情况下得到的。

证明:已知7个铲位同时工作,一班次为8小时,铲车装载时间为5分钟,每辆卡车满载量为154吨,可求出总产量的上界 $=7\times8\times(\frac{60}{5})\times154=103488$(吨)。故猜想正确。证毕。

依据猜想可知,在对车辆小数化整的调优过程中一定存在等待现象。

我们仍然采用饱和因子 P 改进模型。先作出饱和因子 P 和最大总产量的曲线图,如图3所示。

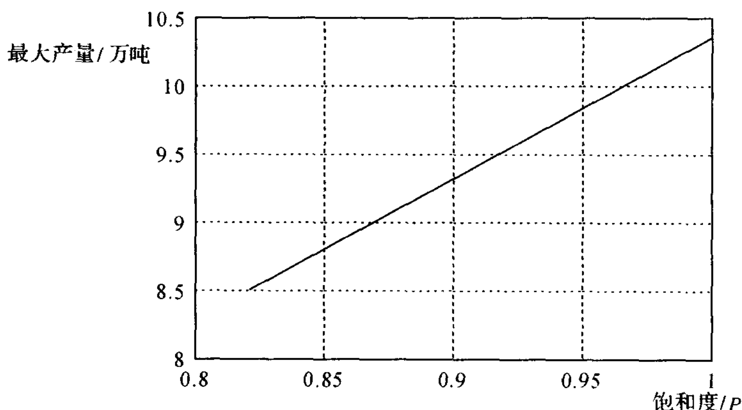


图3

从图像可以看出,饱和因子 P 和最大产量在一定区域内几乎成线性关系,权衡考虑后,此例我们仍设定 $P=0.9$。

同理,将车数辆小数部分调优为整数,并配出各辆车的车次安排如下:

(数字为卡车的编号,括号中的数为该车的运行次数,如6(10)表示第6号车在此路线上运输10次)

P=0.9 时的安排表(单位:辆)

	铲位 1	铲位 2	铲位 3	铲位 4	铲位 8	铲位 9	铲位 10
矿石漏			1.8382		0.9969	0.3223	
倒装场Ⅰ	0.1759	1.7390	0.0809	1.8243			
倒装场Ⅱ			1.1831		1.7051		0.3462
岩石漏	1.8166	0.5899	0.7392	0.6013			
岩场					0.0767	1.8815	1.5232
总产量(万吨/公里)	9.3139	总岩石产量	4.4352	总运输量	12.6953	车数	17.4404

	铲位 1	铲位 2	铲位 3	铲位 4	铲位 8	铲位 9	铲位 10
矿石漏			7(19)+6(16)		16(30)	14(15)	
倒装场Ⅰ	1(5)	3(40)+4(29)	6(3)	17(38)+18(31)			
倒装场Ⅱ			5(20)+4(4)		9(32)+8(4)+10(19)		15(16)
岩石漏	2(45)+1(37)	8(18)	14(26)	15(19)			
岩场					13(2)	12(39)+13(34)	11(46)+10(25)
总产量(万吨/公里)	9.4248	总岩石产量	4.4814	总运输量	12.8533	车数	18

3.快速算法

模型四

由于矿车调度是一个与生活实际密切相关的事,我们希望模型的求解方法尽可能简单,所以设计求解的快速算法是十分必要的。我们考虑在模型二的基础上,进行模型的简化,得到一个快捷方便的算法。



算法介绍：

第一步：规定一个饱和因子 P（$P=1$ 说明允许铲车不停歇工作），通常取 $P=0.9$。

第二步：以各条线路上的卡车数量为求解量，卡车总运量为目标函数，以饱和因子为约束，应用线性规划解法求解。

第三步：判断铲车数是否满足要求，满足则跳到第四步，否则去掉对流量贡献最少的的一个铲位，回到第二步。

第四步：采取人工调优法（或机调法）进行调优。

此模型比较简单，速度较快，但不能保证求到最优解。不过很适合于非技术人员的一般使用。

六、模型的检验（仿真）

由于上述连续模型对卡车调度的考虑是宏观上的，缺乏微观上的考验，因此我们模拟了现实中的多次随机情况，来观察模型的适应度。我们在程序中将一个循环模拟为现实中的一秒钟，用以下算法得出需要等待的车次，记做 N。显然 N 越小，表明我们的模型在微观上适应性越强。算法流程图如图 4 所示。

经过多次模拟实验，我们得出 N 的平均值约为 19，估算一下所有车的停留次数约为 $8\times60/12\times13=520$，12 表示每辆卡车平均每 12 分钟到达一个停留点。

$N/520$ 约为 3.66%，由此可见误差还是很小的，所以我们的连续模型有较好的适应性。

七、进一步考虑的问题

1.实时系统

我们由目标出发，构造了合理的多目标规划模型，并进行了相应的运输流模型求解，使得结果满足了目标要求。然而在实际生产过程中，要实现运输流模型的结果，还必须进行实时调度。实时调度，即在一个班次内，卡车并不固定地配置给某一铲位（铲车）及卸点，而是根据整个露天矿生产中采装、运输、及排卸各环节的实际状况及调度准则进行随机调配。

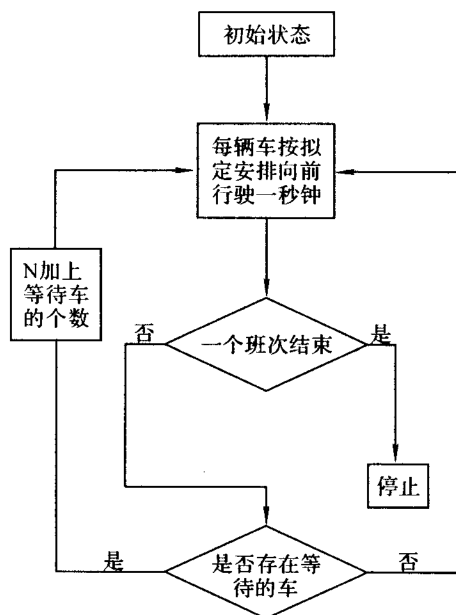


图 4

由于整个运行系统在作业过程中存在着大量随机现象，这就决定了采用实时调度的普遍意义和必要性。我们查阅了专业方面的文献[4]，了解到现有常用的实时调度准则有如下几条：

①最早装车法(EL)：将卡车派往预计能最早得以装车的那台电铲去；

②最大卡车法(MT)：将卡车派往预计其等装时间最小的那台电铲去；

③最大电铲法(MS)：将卡车派往预计电铲等车时间最长的那台电铲；

④最小饱和度法(MSD 和 MSDTC)：MSD 法是将卡车派往具有最小"饱和"程度的电铲，MSDTC 则在此基础上兼顾了卡车运行周期。

对以上各准则在露天矿运输系统模型中进行模拟计算，可得如下结果：



①在车铲比较小的情况下(车铲比为 1，2，3)，EL 和 MT 较优，可得到高于其他准则的产量值;随着车铲比的增大，EL 及 MT 的优势减弱，而 MSD 和 MS 的优势增强，且 MSD 和 MSDTC 的值趋于一致，达到饱和后两者相同;

②在模拟实验中发现，各机动配车调度准则的调度效果在系统中卡车趋于饱和而又未达到饱和时最好，调度潜力发挥最大，特别是 MSD 准则。

③每一准则都有其优势范围，没有哪一种准则在任何情况下都处于绝对领先的地位。

由准则模拟与分析可见，不同的实时调度准则有其各自的适用范围，在不同的矿场条件下，准则的调度效果差异较大。为充分发挥各实时调度准则的优势，达到优化的调度效果,我们提出了综合优化调度准则进行优化调度。

综合调度准则，即根据运输系统条件的变化，确定具体准则的选用，从而使系统始终处在最优准则的调度之下。

通过模拟试验发现,在车铲比为 1 时采用 EL 准则,车铲比为 2～3 时用 M T 准则，车铲比为 4～5 时用 MSDTC 或 MT,车铲比继续增加时用 MS 或 MSD。在已建的模型中增加综合调度模块，进行模拟研究，由模拟结果可以看出，综合调度准则的运用要优于单一准则的运用，发挥了各准则的优势，可以使系统在各种不同的条件下均能取得最优调度效果。

2. 卸点可变化

模型为了实现方便,假定卸点不移动。但从长远看,卸点是可以移动的。实际上用我们的模型二、模型三是完全可以解决这一问题的。由于模型求的是一个班次内的调度方式,只需在每个班次开始前动态地输入 L_{ij} 值,建立一张新的运输流图,再如上求解模型。

3. 卡车在各路线上速度不相等

为了求解的方便,问题中设定卡车在各路线上车速相等。而在实际生产中,由于空重车载重不同等因素,卡车的车速不一定都相同。我们在

现有模型基础上做一定的推广，假设空重车用车重标记 Ki 区分开。分别作出空车和重车的调度计划，将两个调度计划在满足流量连续性条件下统一起来，即相当于在某节点处流量的流入和输出相等。由于时间限制，我们没有深入求解，但思想仍是可取的。

4. 着眼于长期的规划问题

本问题只要求我们给出一个班次内的最优生产和调度计划。这是一个短期目标优化的问题，而一个完整合理的计划还应包含一个着眼于长期的规划。按照日、周排产的短期计划应遵照长期规划所赋予的约束条件。由于题目限制，我们也没有深入研究这个问题，但可以作为今后继续努力的方向。

5. 多种车型的情况

本实例中，矿场只有一种型号的卡车。我们的车流模型可以很好地求解。可以证明，当矿场有多种型号的卡车时，即卡车的满载量和车速不全相同时，我们的模型仍具有很好的适用性。

证明 实际上，我们建立的车流模型并不是把卡车看作一辆辆的实际运动物，而是采用水流模型的思想，把运输问题平均化处理。表征车辆的参量不再是其满载量和车速，而是其流密度 $\frac{q}{\Delta T_{ij}}$，这样多型号卡车问题的处理就和单一型号卡车处理方法一致了，不过“车流管道”是由不同种流密度的流量汇合而成的。

6. 铲车发生故障的考虑

在实际生产中，设备故障是一个较大的影响影素。我们应该进一步考虑在调度系统中如何反映与处理故障，以及其对调度有何影响。

在模拟模型的基础上按生产的实际情况，以如下方式处理设备故障

①对于铲车故障，当已确定卡车配置时，若故障时间大于某一设定值(如 10 分钟)时，则原来配给该铲车的卡车，按到达该铲位的先后次序，均匀分配到其他铲车采装，故障排除后仍配给原铲车；若故障时间小于某一设定值(如 10 分钟)，则相应卡车应在该铲车旁等待。

②对于卡车故障，当已确定卡车配置时，则该卡车停止作业，其当前



时刻加上故障排除时段，此时该卡车恢复生产状态，仍配给原来的铲车。

八、模型的讨论

1. 灵敏度分析

在模型的求解和改进部分，改变饱和因子 P 的取值，作出饱和因子 P 与目标 f 的关系图，并用计算机算出对于不同 P 值时的配车方案。可以分析出，我们的模型解法是比较稳定的，其解不会因 P 的改变而发生很大的变化。

2. 模型稳定性分析

我们一直很强调模型的稳定性。这里的稳定性指的是模型的合理程度分析，引入饱和因子 P 就是考虑稳定性的结果。

3. 模型的评价与推广

用运输流模型求解，避免了对随机因素的讨论，是该模型的优势。同时饱和因子的调整使模型更具实际意义。然而这种平均化的假设可能会对运输成本带来不利的影响。

该模型应用范围广泛，若再利用控制理论[5]对流量模型和实时系统加以结合，必将取得更佳的结果。

九、参考文献

[1]杨启帆.方道元，数学建模，杭州：浙江大学出版社，1999

[2]胡良剑.丁晓东.孙晓君，数学实验使用 MATLAB，上海：上海科学技术出版社，2001

[3]钱颂迪等，运筹学，北京：清华大学出版社，1990

[4]苏靖.刘胜富.张幼蒂，计算机控制卡车多目标综合优化调度的研究，中国矿业，V 6，N3，1997

[5]姜启源，数学模型，北京：高等教育出版社，1993

附录一 模型二的灵敏性分析数据

$P=1$ 时

	铲位1	铲位2	铲位3	铲位4	铲位8	铲位9	铲位10
矿石漏		0.8183			1.8199		0.3031
倒装场Ⅰ		1.0489		1.1583			
倒装场Ⅱ		0.7194					1.4865
岩石漏	1.8166		1.1821				
岩场						1.8183	0.3227
总运量	8.4829(万吨公里)			卡车数	12.4941		

$P=0.95$ 时

	铲位1	铲位2	铲位3	铲位4	铲位8	铲位9	铲位10
矿石漏		0.8183			1.8199		0.3031
倒装场Ⅰ		1.0489		1.1583			
倒装场Ⅱ		0.7194					1.4865
岩石漏	1.8166		1.1821				
岩场						1.9437	0.2182
总运量	8.5193(万吨公里)			卡车数	12.5150		

$P=0.9$ 时

	铲位1	铲位2	铲位3	铲位4	铲位8	铲位9	铲位10
矿石漏		0.8183			1.8199		0.3031
倒装场Ⅰ		1.0489		1.1583			
倒装场Ⅱ		0.7194					1.4865
岩石漏	1.8166		1.1821				
岩场						2.0692	0.1138
总运量	8.5554(万吨公里)			卡车数	12.5361		

从以上三组数据中可以看出，当饱和因子 P 改变时，各路线上的卡车分配变化不大，除了个别较特殊的路线上卡车分配有了很小的波动。由此可知我们的模型是比较稳定的，因此各工作点可以腾出很多空闲来

处理临时特殊情况,甚至包括车辆事故之类。

点 评

“露天矿生产的车辆安排”涉及如何用数学方法降低工业生产成本,提高企业经济效益,其思想可应用于物流调度这一目前各方注目的热点问题。该题对参赛学生的要求是多方面的,其中突出的是各种数学规划的概念和求解方法。问题求解包括两个步骤,一是求最佳物流方案,即确定从不同铲位到不同卸点的车次,要求选手具有根据问题描述写出数学规划的能力,重点是全面恰当地列出所有约束,难点是铲位的选择如何在规划中体现,本文采用了引入 0-1 变量的方法。二是根据最佳物流确定派车方案,即确定每辆卡车的运输路线,难点是联合派车方案的确定,要求选手具有用启发式算法求解 NP-困难问题的初步思想。求解过程又需要选手了解整数规划与其松弛线性规划之间的关系,多目标规划的处理方法等等。

在赛后我们了解到,一些参赛队由于未接触过 LINDO 等软件,在用 C 语言编程或 Maple 等强于形式演算的软件求解大型规划时遇到了困难。一些队在变多目标规划为单目标规划时权重过于随意,从而得到的解不合要求。也有些队第二问中岩石产量超过 6 万吨,事实上,简单的计算可知一个班次中岩石产量上限为 49280 吨。如何在建模过程中发现问题和及时调整,避免顾此失彼、少走弯路,也是在参赛时需要注意的。

这篇论文是获奖论文中较为出色的一篇,除了很好地解决了问题,得到了非常接近最优解的方案之外,文章还进行了仿真试验,提出了许多可进一步研究的问题。这些都反映了参赛队员平时良好的知识储备和三天的艰苦努力。

本文获 2003 年全国大学生数学建模竞赛全国一等奖。



奥运会临时超市网点设计方案

姚 烁 陆 熙 张 宁

摘 要

为了使模型的建立更加准确,我们对数据的统计考虑到了很多方面:性别对消费的影响,各出行方式所占的比例,各餐饮方式所占的比例,被访观众的年龄分布,各年龄段的消费能力的期望等。对第二问人流量的计算,我们提出假设:观众出行及餐饮时只能走人行横道和穿越商区,不能穿越草坪。这样,对出行路径的选择给出了约束。运用C语言编写了计算最短路径的程序,以便快速寻找到出行路径。对于人流量的计算,我们建立了流量矩阵,并运用matlab进行矩阵运算求解。对于第三问,在销售额和需求量的计算上,我们用了预期需求量计算模型:$D=pC$(D:需求量,p:观众的购物可能性,C:人数),并对p定义为:

$$p_{ij}=\frac{i\text{ 商区在 }j\text{ 商区到目标路口最短路径中出现的次数}}{j\text{ 商区到目标路口最短路径中出现的商区数}}$$

还提出了预期购物人次叠加的思路,并对此模型中p_{ij}的定义用huff模型进行了改进。

在计算销售量时,提出计算模型:$Q=DE$(E:人均消费期望值,Q:销售额)。在解决MS的分配问题上,我们根据基本要求建立了有4个约束条件的线性规划,用LINGO进行求解。此外,我们还引入零售管理学中的饱和理论,作为线性规划约束条件的改进。

在对模型科学性的分析上,由于题目所给出的调查数据样本太少,我们对最后的结果进行了统计假设检验,证明了对于调查总体,我们的结果仍较合理。而且我们还上网查阅了相关资料,参照雅典奥运会的销售情



况,从人口比例,人均消费等方面与北京进行比较评估,再对照我们得到的结果,进行了合理性分析。

流量统计结果:

	百分比		百分比		百分比
*A*1	0.069951	*B*1	0.037315	*C*1	0.025521
*A*2	0.043462	*B*2	0.031755	*C*2	0.028267
*A*3	0.049023	*B*3	0.046536	*C*3	0.025521
*A*4	0.054583	*B*4	0.031755	*C*4	0.057311
*A*5	0.060143	*B*5	0.037315		
*A*6	0.119992	*B*6	0.074337		
*A*7	0.060143				
*A*8	0.054583				
*A*9	0.049023				
*A*10	0.043462				

一、问题的重述

北京奥运会期间,在比赛主场馆的周边地区需要建设有小型商亭构建的临时商业网点,称为迷你超市(记做 MS)网。在地点、大小类型和总量方面有三个基本要求:满足奥运会期间的购物需求、分布基本均衡和商业上的赢利。

问题 1:根据附录中给出的问卷调查数据,找出观众在出行、用餐和购物等方面所反映的规律。

问题 2:假定奥运期间(指某一天)每位观众平均出行两次,一次进出场馆,一次餐饮,出行采取最短路径。依据 1 的结果,测算图 2 中 20 个商区的人流量分布(用百分比表示)。

问题 3:如果有两种大小不同规模的 MS 类型可供选择,给出图 2 中 20 个商区内 MS 网点的设计方案(即每个商区内不同类型 MS 的个数),以满足上述三个基本要求。

问题 4:阐明你的方法的科学性,并说明你的结果是贴近实际的。



二、基本假设

(1)观众出行及餐饮时只能走人行横道和穿越商区,不能穿越草坪。

(2)观众进出场馆是指早入晚出,餐饮是指中间出去用餐而后回到场内继续观看比赛。

(3)对于进出场馆的那次出行,观众必须乘车,入场时要从车站步行到场馆并由场馆步行到同一车站。例如某观众早上从公交(东西)下车步行到场馆,那么晚上就必须从场馆步行至公交(东西)再乘车回家,而不能选择到公交(南北)或其他车站。

(4)观众出去餐饮及归来都是步行,不考虑乘车。

(5)对于假设(3)的步行路程要符合路径最短原则。

(6)观众只有在出行时才会去商区购物,而不会单独去购物。

(7)当观众到达某一目的地有两种以上相同路程的路径时,假设走每条路的概率相同。

(8)假设第一问中反映的规律对于三个比赛场馆的观众是同样适合的。

(9)商区的面积足够提供两种 MS 的建构用地。

(10)同一商区的客流量及商业市场饱和系数 IRS 是相同的。

三、参数说明

T_j:第 j 种出行方式或餐饮方式;

$$r_i=\begin{cases}1 & \text{第 } i \text{ 个商区经过第 } j \text{ 个商区}\\0 & \text{第 } i \text{ 个商区不经过第 } j \text{ 个商区}\end{cases};$$

$C_{总j}$:经过 j 商区的总流量;

k_1:建造每个大规模的 MS 的投入成本;

k_2:建造每个小规模的 MS 的投入成本;

S:灵敏度;

x_1:某个商区所开大规模 MS 的数量;

x_2:某个商区所开小规模 MS 的数量;



Q_i:第 i 商区的预期销售额;

D_i:第 i 商区的预期需求量;

α:每个大规模 MS 的日投入成本;

β:每个小规模 MS 的日投入成本;

δ:每个大规模 MS 的日承受需求上限;

ε:每个小规模 MS 的日承受需求上限。

四、模型分析及求解

通过对题目所给数据的分析,找到观众出行的一些规律,从而为分析第二问的流量及第三问的商店选址提供依据。在对数据的分析上,主要集中分析各种出行、餐饮、消费方式所占比例的特点,找到相应的比例期望值。

1.观众在出行、用餐和购物等方面所反映的规律

(1)性别影响消费规律

在被调查的 3500 名观众中,男性与女性的比例为 52.17% 和 47.83%,其中男性比女性高出 4.43 个百分点,由于受外地来看奥运会的观众中男性居多等因数的影响,这个数据不足以反驳“女人天生爱购物”,但是在确定商品经营范围和档次等问题时,也应适当考虑男性观众的审美情趣及消费特点。因此,我们可以说性别不会导致消费差异。

(2)出行规律

公交、出租和地铁是出行的主要工具,累计占了总和的 90% 以上。

而观众对公交,出租和地铁的偏好基本相近。由数据统计所得出的表 1 与图 1 如下:

表 1 选择各种交通方式的人数及比例

	公交(南北)	公交(东西)	出租	私车	地铁(东)	地铁(西)	总人数
人数	613	598	680	308	645	657	3500
比例	0.175142857	0.17085714	0.194286	0.088	0.184286	0.187714	1

这样,我们得到了 6 种出行方式的期望值。通过 MATLAB 软件编

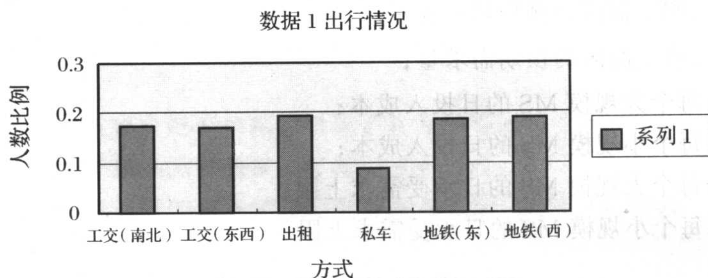


图 1　各种出行方式人数比例

程，对数据的 3500 个样本再次进行随机抽样，抽取样本容量为 1000，抽取了 50 次，再用 MATLAB 对所得的期望值进行检验，置信系数为 0.05，模拟所得结果是假设成立。

(3)用餐规律

大部分人选择了西餐，比例为 50%以上，其余两种所占比例基本相当。对用餐我们用同样的统计分析方法，得到以下结果：

表 2　选择各种餐饮方式的人数及比例

	中餐	西餐	商场	总数
人数	724	1672	804	3200
比例	0.22625	0.5225	0.25125	1

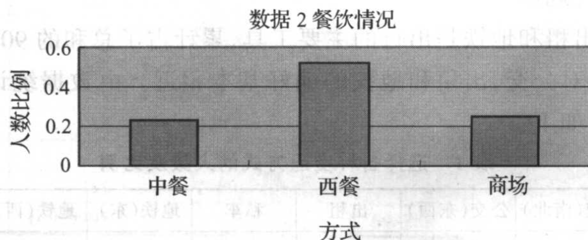


图 2　选择各种餐饮方式的人数比例

用同样方法检验了比例期望值，假设也成立。根据表 2，大部分人倾向于西餐，我们认为这是由于西餐的快捷，可以为观众节省更多的时间来



观赏比赛。

(4)我们还分析了各种出行、餐饮方式中年龄(4档)的构成情况:

出行:乘坐公交,出租,私车,地铁的人群比例中最高的均是是第二年龄档(20~30岁),且均在50%左右,这也反映了观看比赛的主流人群的年龄段。

表3 各种出行中年龄(4档)的构成情况

	公交(南北)		公交(东西)		出租车		私车	
	人数	比例	人数	比例	人数	比例	人数	比例
一	95	0.154976	57	0.095477	60	0.088235	41	0.133117
二	360	0.587276	289	0.484087	400	0.588235	184	0.597403
三	92	0.150082	140	0.234506	153	0.225	64	0.207792
四	66	0.107667	111	0.18593	67	0.098529	19	0.061688
总数	613	1	597	1	680	1	308	1

	地铁(东)		地铁(西)	
	人数	比例	人数	比例
一	67	0.103876	62	0.094368
二	388	0.60155	408	0.621005
三	134	0.207752	128	0.194825
四	56	0.086822	59	0.089802
总数	645	1	657	1

餐饮:第一档偏向于西餐和商场,且差别不大。

第二档人群明显偏好于西餐。

第三档偏向于中餐和西餐,且比例相仿。

第四档对餐饮的偏好并不明显,三种选择比例相近,略偏向于中餐。

表4 餐饮方式中年龄(4档)的构成情况

	中餐		西餐		餐饮	
	人数	比例	人数	比例	人数	比例
一	39	0.049808	181	0.09853	162	0.184091
二	323	0.412516	1258	0.684812	448	0.509091
三	269	0.34355	294	0.160044	148	0.168182
四	152	0.194125	104	0.056614	122	0.138636
总数	783	1	1837	1	880	1

这样我们直观的得到了各种出行方式和餐饮方式中的年龄构成。当然，我们对结果同样用 MATLAB 进行了假设检验。

(5)购物

20 岁以下观众的消费额主要集中在第一档和第二档；

20～30 岁观众的消费额主要集中在第三档；

30～50 岁观众的消费额主要集中在第三档，但第一档和第二档也占不小比例；

50 岁以上观众的消费额主要集中在第一档和第二档

对购物我们主要想得出各个年龄段的消费情况，借此得出各年龄段的消费能力。

表 5　各个年龄段的消费情况

年龄段 / 消费段	一			二		
	人数	比例	金额	人数	比例	金额
一	133	0.348168	0	229	0.112863	0
二	160	0.418848	100	324	0.159685	100
三	62	0.162304	200	1174	0.57861	200
四	16	0.041885	300	260	0.128142	300
五	7	0.018325	400	22	0.010843	400
六	4	0.010471	500	20	0.009857	500
总数	382	1	99.47644	2029	1	179.3987
消费期望值	99.47644			179.3987		
年龄段 / 消费段	三			四		
	人数	比例	金额	人数	比例	金额
一	126	0.177215	0	195	0.515873	0
二	191	0.268636	100	158	0.417989	100
三	338	0.475387	200	16	0.042328	200
四	33	0.046414	300	4	0.010582	300
五	15	0.021097	400	3	0.007937	400
六	8	0.011252	500	2	0.005291	500
总数	711	1	149.9297	378	1	59.25926
消费期望值	149.9297			59.25926		

产生这样的结果是合理的，因为 20 岁以下的观众属于消费前卫群体，他们对新潮消费品兴趣浓厚，推崇名牌，追逐时尚，强调个性，有强烈

的购买欲望，但购买能力受收入情况的影响；而 50 岁以上的观众由于购买欲望不很强，所以购买力也较低。对于 20～50 岁的观众来说，属于普及推动群体，即某种商品流行潮流已经形成才加入的消费者，这类群体收入稳定，消费具有较强的理性色彩，因此成为第一大消费群体。

然后我们得到了被访观众的年龄分布：

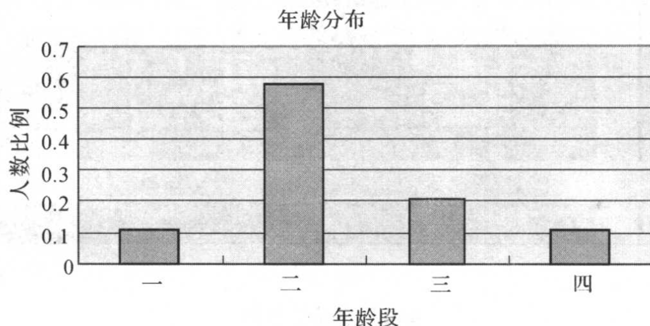


图 3　被访观众的年龄分布

2. 测算 20 个商区的人流量分布

(1)网格分割

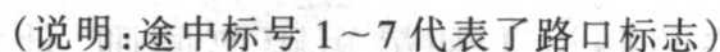
这个模型有一个重要前提(假设)，就是我们的假设 1。由于观众出行的时候只能在人行道上或商区内行走以到达各个目标(如工具类和餐饮类)，所以观众在出行时将会选择最短的折线距离。为了说明的方便及搜索最短路径算法的设计，我们将图 4 用网格分割定位，将距离近似化为网格的数目，即简化成图 5。

对照图 1，为方便说明，我们以 A9 看台为例，假设该看台的观众在出行时要到出租车站，按照假设 1 和最短距离原则，他将按照下面的路线行进，A9→A10→A1→街口 1→街口 2→出租车站，也就是说 A9 看台的观众除了经过 A9 商区，还要经过 A10 和 A1 商区。又如，如果该看台的观众要去公交西，根据距离最短原则，他的路径将是 A9→A8→A7→A6→街口 3→公交西，这样我们计算不同商区的人流分布时，就要进行叠加。

这里特别需要说明的是最短距离是建立在假设 1 的基础上的，且当



图 4　建模结构图



（说明：途中标号 1～7 代表了路口标志）

图 5　网格简化图

观众到达某一目的地有两种以上相同路程的路径时，走每条路的概率相同。如图中的 A6 看台的人若要去中餐，他有两条相等路径可以选择，分别为 A6→A5→A4→A3→A2→A1→街口 1→中餐和 A6→A7→A8→A9→A10→A1→街口→中餐，我们认为观众在这种情况下的选择是随机的。



(2)最短路径的选取

在这个模型的最短路径求解部分,我们用了C++编程进行搜索求解,将20个商区,8个车站,3个餐饮,10个街口和拐角作为41个点构成一个图,先用网格将图形分割,利用网格把图形建立成为加权边图形的邻接矩阵,算法的特点是沿人行道和商区进行加权,这是因为观众必须通过这两个路径行走(假设1)。用一个数组记录已查找的数组(visited)和一个一位数组来记录顶点间距离总和的变化(distance)。取原先的距离总和与新加入顶点后总和的最小值作为新的距离总和。这样任意输入两个点,就可以统计出这两点间的最短路径(注:文后所附的C语言源程序从略)。

(3)总流量的求解

我们用函数 $C_{总j}$ 来表示各个商区的总流量(人次/天)

这个过程可用下列公式描述:

$$C_i = \sum_{j=1}^{9} 10000 T_j \quad (9:表示9种出行方式及餐饮方式)$$

$$C_{总j} = \sum_{i=1}^{20} r_i C_i$$

由此公式,运用C语言源程序将最短路径导入矩阵,再用MATLAB矩阵计算得下表6,反映了各商区的流量:

表6 各商区流量

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	
A1	4000	1356	1356	1356	1356	1356	1356	1356	1356	1356	16204
A2	1322	4000	1356	1356	1356	678	0	0	0	0	10068
A3	1322	2644	4000	1356	1356	678	0	0	0	0	11356
A4	1322	2644	2644	4000	1356	678	0	0	0	0	12644
A5	1322	2644	2644	2644	4000	678	0	0	0	0	13932
A6	2644	2644	2644	2644	2644	4000	2644	2644	2644	2644	27796
A7	1322	0	0	0	0	678	4000	2644	2644	2644	13932
A8	1322	0	0	0	0	678	1356	4000	2644	2644	12644
A9	1322	0	0	0	0	678	1356	1356	4000	2644	11356
A10	1322	0	0	0	0	678	1356	1356	1356	4000	10068
	B1	B2	B3	B4	B5	B6					
B1	4000	2644	1322	0	0	678	8644				

续表

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	
B2	1356	4000	1322	0	0	678	7356				
B3	1356	1356	4000	1356	1356	1356	10780				
B4	0	0	1322	4000	1356	678	7356				
B5	0	0	1322	2644	4000	678	8644				
B6	2644	2644	2644	2644	2644	4000	17220				
	C1	C2	C3	C4							
C1	4000	1546	0	366	5912						
C2	908	4000	908	732	6548						
C3	0	1546	4000	366	5912						
C4	3092	3092	3092	4000	13276						

表 6 是为计算建立的流量矩阵，对每一种出行方式及餐饮方式均建立一个对应的流量矩阵。行代表到达的商区，列代表来自的商区，如数据 $A1-A2$ 代表从 $A2$ 到 $A1$ 有 1356 人。而每一行的最后一个数就是累加的总流量。

下面是各商区流量的百分比数据(表 7)和相应的柱形图(图 6)：

表 7　各商区流量的百分比数

	百分比		百分比		百分比
A1	0.069951	B1	0.037315	C1	0.25521
A2	0.043462	B2	0.031755	C2	0.028267
A3	0.049023	B3	0.046536	C3	0.025521
A4	0.054583	B4	0.031755	C4	0.057311
A5	0.060143	B5	0.037315		
A6	0.119992	B6	0.074337		
A7	0.060143				
A8	0.054583				
A9	0.049023				
A10	0.043462				

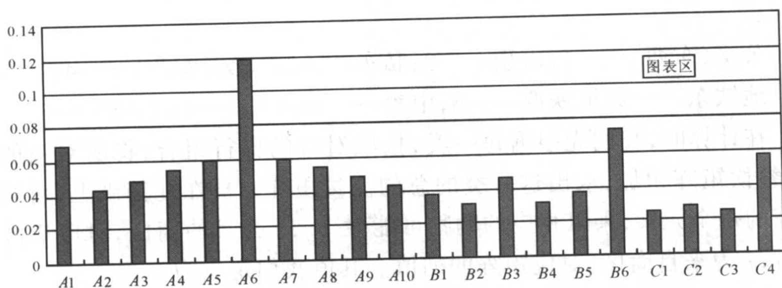


图 6　各商区流量的百分比数据柱形图

3. 关于各商区大小 MS 分布的设计方案

我们的规划模型是让各个商区中 MS 建造投入的总成本最低。

(1)各商区需求量的求解

关于这个模型,我们首先要考虑各个商区的期望需求量是多少。需求量期望值可用以下公式表示:

$D=p\times c$。(D:需求量期望 p:观众的购物可能性 c:潜在顾客人数)

由于每人要出行两次,一次是进出场馆(坐交通工具),一次去为餐饮。无论走的是哪条路线,都必须经过路口对应的商区(比如 A1,B3 等)。对各个大的商区,按观众出行所经过的路口对出行及餐饮方式的组合进行分类:

A: $A1A1=[1a,2a,3a]$

$A1A6=[4a,5a,6a,1b,2b,3b,1c,2c,3c]$

$A6A6=[4b,5b,6b,4c,5c,6c]$

B: $B3B3=[1a,2a,3a]$

$B3B6=[4a,5a,6a,1b,2b,3b,1c,2c,3c]$

$B6B6=[4b,5b,6b,4c,5c,6c]$

C: $C2C4=[1a,2a,3a,1b,2b,3b,1c,2c,3c]$

$C4C4=[4a,4b,4c,5a,5b,5c,6a,6b,6c]$

注:A,B,C 分别代表三个区,A1A1 代表出行和餐饮全部从 A1 区所对路口出去,A1A6 代表出行和餐饮分别从 A1A6 所对路口出去,其他与此相同。1－6 代表 6 种出行方式,$a-c$ 代表 3 种餐饮方式。分别

对应：

公交（东西）——1；出租——2；私车——3；公交（南北）——4；

地铁东——5；地铁西——6；中餐——a；西餐——b；商场——c。

在计算时，我们先根据每一类，按照对应的出行组合，将所对应的出行、餐饮矩阵相加，求出这一类的总的流量矩阵。再在此基础上，对相应矩阵的每个元素，乘以相应的购物可能性 p_{ij}，i 为目的商区，对应矩阵的行值，j 为来自商区，对应矩阵的列值。其值我们定义为

$$p_{ij}=\frac{i\text{ 商区在 }j\text{ 商区到目标路口最短路径中出现的次数}}{j\text{ 商区到目标路口最短路径中出现的商区数}}$$

如在 A 类中的 $A1$，代表出行只能从 $A1$ 路口出去，则此时 p_{12} 代表从 $A2$ 出去经过 $A1$ 商区的人的购物可能性，按照定义和路线图，我们可以知道其值为 1/2。按照这样的公式，借助计算机进行矩阵运算，我们得到了各商区的需求量矩阵（表 8）：

表 8　各商区的需求量矩阵表

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	总需求
A1	2968.6	788.71	601.05	507.21	450.91	639.1	450.91	507.21	601.05	788.71	8303.46
A2	490.29	2519.6	601.05	507.21	450.91	206.69	0	0	0	0	4775.75
A3	490.29	1094.6	2502.9	507.21	450.91	206.69	0	0	0	0	5252.6
A4	490.29	1094.6	1265.6	2694.1	450.91	206.69	0	0	0	0	6202.19
A5	490.29	1094.6	1265.6	1550.6	3207.8	206.69	0	0	0	0	7815.58
A6	980.57	1094.6	1265.6	1550.6	2120.6	4880.2	2120.6	1550.6	1265.6	1094.6	17923.57
A7	490.29	0	0	0	0	206.69	3207.8	1550.6	1265.6	1094.6	7815.58
A8	490.29	0	0	0	0	206.69	450.91	2694.1	1265.6	1094.6	6202.19
A9	490.29	0	0	0	0	206.69	450.91	507.21	2502.9	1094.6	5252.6
A10	490.29	0	0	0	0	206.69	450.91	507.21	601.04	2519.6	4775.74

	B1	B2	B3	B4	B5	B6	总需求
B1	3866.9	1714.8	714.9	0	0	298.75	6595.35
B2	691.33	3484.6	714.9	0	0	298.75	5189.58
B3	691.33	879	3762.6	879	691.33	597.5	7500.76
B4	0	0	714.9	3483.6	691.33	298.75	5189.58
B5	0	0	714.9	1714.8	3866.9	298.75	6595.35
B6	2284.8	1714.8	1429.8	1714.8	2284.8	5483.1	14912.1

	C1	C2	C3	C4	总需求
C1	4500	1148	0	137.25	5785.25
C2	340.5	3727	340.5	274.5	4682.5
C3	0	1148	4500	137.25	5785.25
C4	3069	2296	3069	6819	15253

(2)各商区预计销售额的求解

利用下面公式计算每个商区的预计销售额:

$$Q = D \times E$$

由前面统计出的各年龄段消费能力的预期值与各年龄的分布,可以得出消费的预期值:

$$E = 151.7143$$

这样,得出各商区的预期销售额图表(表9)和柱形图(图7):

表9 各商区的预期销售额

	销售额		销售额		销售额
A1	1259754	B1	1000609	C1	877705.2
A2	724549.6	B2	787333.5	C2	710402.2
A3	796894.5	B3	1137973	C3	877705.2
A4	940960.9	B4	787333.5	C4	2314098
A5	1185735	B5	1000609		
A6	2719262	B6	2262379		
A7	1185735				
A8	940960.9				
A9	796894.5				
A10	724548.1				

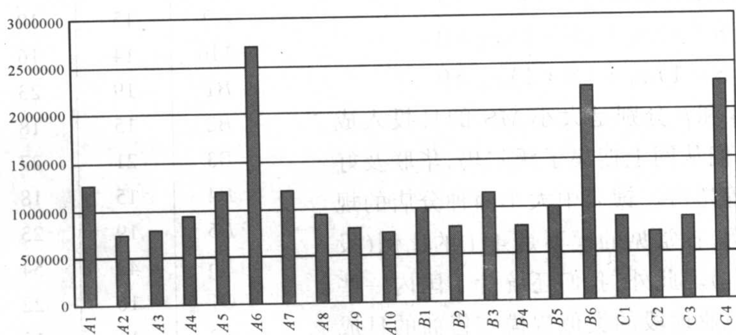


图7 各商区的预期销售额柱形图

(3)各商区MS商店设立个数的确定:

对于两种规模 MS 商店的选址，要满足购物需求、分布基本均衡和商业上的赢利。因此我们可建立数学模型如下：

k_1, k_2 分别表示建造大小不同规模的 MS 的投入成本，这里假设 $k_1/k_2 = \alpha/\beta$，且等于店铺的面积比。S 是灵敏度，用来约束大小 MS 的分布均衡。

于是，可以将问题转化为有三个约束条件的线性规划模型

目标函数：　$\min(k_1x_1 + k_2x_2)$

s.t.

$(S+1)x_1 + (S-1)x_2 > 0$

$(S-1)x_1 + (S+1)x_2 > 0$

$\alpha x_1 + \beta x_2 \leqslant Q_i$——满足商业盈利

$\delta x_1 + \epsilon x_2 \geqslant D_i$——满足购物需求

$x_1 \geqslant 0, x_2 \geqslant 0$

x_1：一个商区所开大规模 MS 的数量

x_2：一个商区所开小规模 MS 的数量

其中，$\left|\dfrac{x_1 - x_2}{x_1 + x_2}\right| < S$ 为满足分布基本均衡（相对）。

化为线性约束条件：

$(S+1)x_1 + (S-1)x_2 > 0$

$(S-1)x_1 + (S+1)x_2 > 0$

α 和 β 分别是大小 MS 的日投入成本，我们从网上搜集了沃尔玛，华联及好又多等超市连锁店中大小两种分店的规模比例，较常见的模型是 4:1 的比例（投资成本）。此外，我们还统计了国内一些繁华商业地段的类似规模零售商的日成本投入，但考虑到奥运会将为零售业带来巨额利益，日投入成本肯定要偏高。

表 10　各商区在最优情况下大小规模 MS 的安排数量

	x_1	x_2
A1	24	29
A2	14	16
A3	15	19
A4	18	21
A5	22	28
A6	50	65
A7	22	28
A8	18	21
A9	15	19
A10	14	16
B1	19	23
B2	15	18
B3	21	27
B4	15	18
B5	19	23
B6	42	54
C1	16	22
C2	14	15
C3	16	22
C4	43	55



我们查到,北京王府井的工美奥运特许商店是一个已开售的特许商品零售店,它的日投入资本在3～6万元,在高峰期的销售额可达10万元。因此我们不妨可把δ和ε相应地设为4.0万元/天和1.0万元/天。同理δ和ε也是我们经过对网上数据信息进行处理得出的统计值。分别取为$\delta=220$(人/天),$\varepsilon=107$(人/天)

有了这4个系数,将各商区的相应数据代入,使用LINGO软件求解线性规划,十分方便就得出结论如表10所示(针对该组参数):

实际参数选取:

$S=0.02$

4.模型的检验

(1)科学性

由于在建立和计算数据时,我们所用到的数据是从题目已知的数据表中查到的。但数据表只提供了3份调查报告,这在统计上是不准确的。下面,我们用Z检验方法对用题目提供的数据代入模型所求得的解进行假设检验,看这小样本所得到的结果是否能说明:如将总体代入我们的模型,3个约束条件仍然满足。

将上面计算出的商店的个数代入到约束条件中,对于盈利的条件,计算每个商区的$\alpha x_1+\beta x_2-Q_i$值,作为v_1。我们规定$v_0=0$。要实现盈利,就要检验不等式$v_1<v_0$是否成立。利用MATLAB的ZTEST函数进行检验,即统计学中的Z检验,输出结果为假设成立。这样,就表示我们所得到的这个结果能使总体满足约束条件。

下面是MATLAB的检验结果:

期望 = −1.057207

标准差 = 1.403877

检验z值 = −3.367799

检验结果 = 1.000000

(注:检验结果为1表示备检接受假设。)

对于需求条件的检验与上相同,结果也是假设成立。

(2)合理性

①从图6和图7中的柱形图可直观地看到,预期销售值的变化趋势

与人流分布的趋势在总体上是一致的，这是因为预期销售值与人流分布是正向相关的。但通过仔细观察，还是有差别的。例如 $B1$ 区，在人流分布图上要低于 $A10$ 和 $B2$ 商区，但销售预期值却较二者更高。这也是符合实际的，因为顾客的购物欲望也是影响预期销售值的重要因素，而经过相同地区的人流量对商品购买欲望可以是不同的，这在问题三的解答中我们有很好的解释。

②第三问中规划的结果保证了各个商区大小 MS 的比例(比值)较为均衡，见图 8。

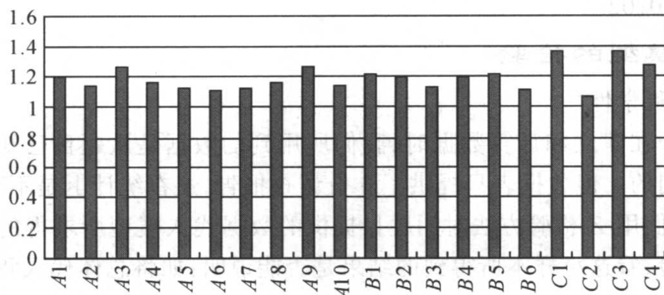


图 8　各商区大小 MS 比例的柱形图

因为北京奥委会对体育场馆周围的特许商品经销商的统一规划与管理是十分严格的(包括价格的统一等)，因此这个结果是符合实际的。

③20 个商区 MS 的总和为 978 个，如果把它看作是竞赛区直销部分，根据北京奥委会提供的数据，整个北京市的总特许销售点约为竞赛区销售点的 10 倍，就是说北京约有 10000 个特许商品零售点。由实际数据信息：北京人均年收入是 2.4 万人民币，城市人口 800 万，雅典的年均年收入 0.6 万欧元，约合人民币 8.34 万元，城市人口 300 万，奥运期间共设 7600 个零售便利店，而据专家预测北京奥运的零售总额将与雅典持平甚至超越。假设消费水平与收入水平成正比，则北京的总零售便利店将大约是 $\frac{8.34\times 300}{2.4\times 800}\times 7600=990$ 个，这与我们的结果大体是吻合的。



五、模型的改进

(1)我们在第三问中统计及计算各个商区的预期需求时并没有按照年龄规律和偏好来计算,而实际上年龄的差异直接导致出行方式的差异,也就是说导致出行路径的差异,最终导致各商业区流量的差异。只是这种算法十分繁琐,我们只是提了出来,但这无疑是更精确的算法。

(2)从上述结果,我们可以看到,商店的分布呈一定的规律性和周期性。由于开始定义购物可能性时,定义过于简单,造成在很多商区的购物可能性是一样的。这样,流量所表现的规律性就传递到了需求量和预期销售额中。下面,我们对 p_{ij} 重新定义:

引用 huff 模型[7]中对 p_{ij} 的定义:

$$p_{ij} = \frac{S_j F(d_{ij})}{\sum_{j=1}^{m} S_j F(d_{ij})}$$

其中:S_j 为 j 商店的规模,$F(d_{ij}) = d_{ij}^{-\lambda}$,$d_{ij}$ 为 i 商区到 j 商区的距离,λ 为距离衰减系数。这样定义 p_{ij},是因为有一个距离衰减系数的存在,这使得各商区的 p_{ij} 可以不同,其值大小受到距离的影响,这是因为在购物心理学上,观众在出行过程中总倾向于在距离自己看台较近的地方购物,这样引入距离衰减因子以后,模型会显得更加的合理,使商店的分布表现得更加的随机,趋于总体均衡。但由于 λ 的值缺乏统计数据,故我们无法对这一改进的模型进行求解及验证。

(3) 另一个改进是引入零售管理学中的饱和理论[3],商业饱和系数的表达式是:IRS=(H)(RE)/(RF)

式中:IRS 为某地区某类商品的零售饱和系数;

RE 为某地区某一顾客用于购买某类商品的费用支出;

H 为某地区购买某类商品的潜在顾客人数;

RF 为某地区经营同类商品商店的营业面积总数。

利用抽取调查的方式或计算机模拟的方法来预测或评估 IRS 的值,再利用测得的 IRS 值可得出各商区商业建筑面积的数值,把各商区商建面积用于第三问数学规划中的赢利的约束条件上即可求出超市网点的最

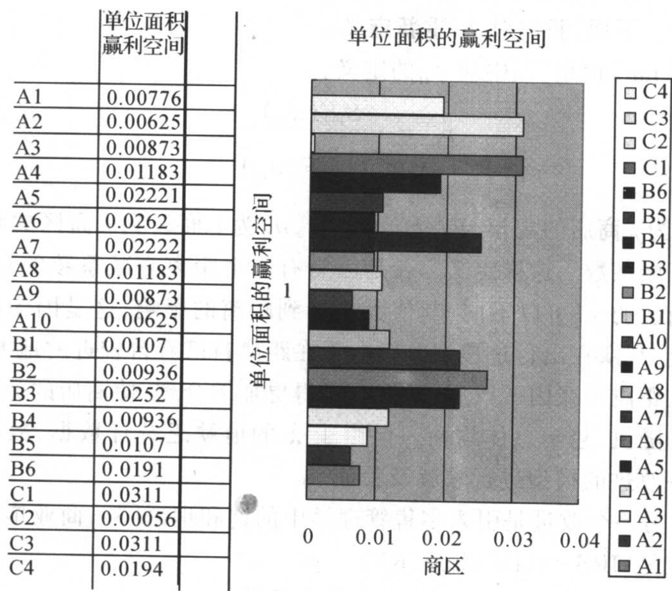
佳布局。

六、模型的应用

我们引入单位面积利润空间(期望)q 的概念,定义单位面积利润空间如下:

q =(销售期望值－日投入成本)/商建面积

以其中一组参数为例来加以说明(见表 10)。根据我们作出的各商区大小 MS 的分布,得出单位面积利润空间的分布如图 9:



	单位面积赢利空间
A1	0.00776
A2	0.00625
A3	0.00873
A4	0.01183
A5	0.02221
A6	0.0262
A7	0.02222
A8	0.01183
A9	0.00873
A10	0.00625
B1	0.0107
B2	0.00936
B3	0.0252
B4	0.00936
B5	0.0107
B6	0.0191
C1	0.0311
C2	0.00056
C3	0.0311
C4	0.0194

图 9

从以上图表分析可知,对商家而言,投资在单位面积利润空间大的商区获利的可能性较大,具有较好的商业前景,如图中 C3,C1,B3,A6。奥组委在上述地区征收的租金和盈利抽取比例应较其他地区稍高一些。



七、参考文献

[1]姜启源、谢金星、叶俊.数学模型.北京:高等教育出版社,2003年8月

[2]王超.零售学.北京:中国对外经济贸易出版社,1999

[3] Barry Berman, Joel R.Evans.零售管理.北京:中国人民大学出版社,2001

[4] 张铮等.MATLAB程序设计与实例应用.北京:中国铁道出版社,2003

[5] 薛毅.数学建模基础.北京:北京工业大学出版社,2004

[6] D.R.安德森,D.J.斯威尼.商用统计问题详解.北京:晓园出版社,1992

[7] 杨丽君、朱华岚、吴建平.基于GIS的零售业商圈分析.遥感技术的应用.18卷3期,2003

点 评

本文作者在分析调查数据的基础上提出了一些比较符合客观实际又便于建立数学模型的10条假设,作者较详尽地分析了奥运会期间观众在出行、用餐和购物方面的各种规律,并根据题目要求建立了优化奥运会期间临时超市网点布局的线性规划模型。作者参考一些大超市的统计数据,设定了模型参数,利用LINGO软件求出结果并对模型的科学性、合理性进行了统计检验。

本题是一个较易掺入主观因素的研究课题,应当尽可能避免主观臆断。从文章中可以看出,作者已注意到这一点,用了相当大的篇幅来分析实际数据,以便尽可能找出符合客观实际的规律。在确定参数时,他们也尽量遵从事实,采用了取自大超市的实际数据。本题虽只是一个竞赛题,参赛学生不可能在短短的3天竞赛时间里真正去作市场调研,但从文章中可以看出,作者是在作实际课题的模拟,有理由相信,如果他们有充足的时间,一定会首先去作充分的市场调查和分析,使自己的研究结果更具有说服力,从而对决策者具有更大的参考价值。

本文获得2004年全国大学生数学建模竞赛全国一等奖

电力市场的输电阻塞管理

哈 阳 周 楠 张 博

摘 要

本文在电力市场交易规则和输电阻塞管理原则的前提下,重点对输电阻塞时如何管理提出了行之有效的方法,给出了出现电力阻塞时各种情况的重新分配方案,并在安全第一的前提下做到了损失最小。首先根据题中的33组数据利用多元线性回归的方法给出各线路有功潮流与各机组出力的近似关系式,并利用显著性分析进行了改进;对于阻塞费用的计算规则,先模仿Arrow不可能性定理证明了完全公平是不可能实现的,然后提出在Pool模式的基础上考虑责任分摊的改进算法,分别给出了Pool和在此基础上的两种阻塞费用的计算法则,并进行了比较和分析;对题中所给分配预案的确定原则给出了最优证明;对最小阻塞费用的计算是建立一个非线性规划模型,并分别通过lingo直接求解,引入决策变量转化成0—1规划,利用分枝定界的思想转化为多个线性规划等多种方法得到实现;对于使用裕度发电的情况给出了多目标线性规划,并提出了更加巧妙的MIN(MAX)函数规划,增大了安全性且对其优越性做了分析。

当预测负荷为982.4MW时,清算价为303元/MWh,调整时不用使用安全裕度,Pool模式下的最小阻塞费用为4694元,考虑责任分摊后减少为4166元(减少部分由发电方按责任承担)。当负荷需求为1052.8MW时,清算价356元/MWh,必须使用安全裕度输电,在首先保证安全的条件下得到最小阻塞费用为1420元,按照改进的、更安全的分



配则为 2681 元。

最后给出了拉闸限电分析，灵敏度分析以及算法稳定性分析，很好的证明了算法的可行性。

一、问题重述

进入新世纪以来，电力行业也逐渐进入市场化模式并得到长足的发展，电已经成为了一种商品。然而，由于电作为商品的特殊性（需求随时间变化很大，且难以存储），加之我国电力市场还处在初期发展阶段，现在还只能采取交易与调度的一体化模式。

电网公司在组织交易、调度和配送时，要首先把安全问题放在第一位考虑，在此前提下，要尽量使得购电的总费用最小。

实际操作时，交易一调度中心先对下个时段的用电负荷做一个预测，再根据各个发电机组的出力（发电功率）和在各个发电段容量的报价决定选择一组出力分配方案以满足总费用最小的条件。其中机组的报价可以为负，表示发电厂为了维持机组的运转而愿意付给网方费用以减少更大的损失。在实际生产中，当火力发电机输出的功率很低时，要向炉子中注油以确保火不熄灭，此举将大大提高成本。

选择机组时，采取报价由低到高的原则，当它们之和等于预测值时，则取最后一个被选入的价格为清算价，该时段所有机组的出力均按清算价结算。我们可以用一个队列的数据结构来表示。先按价格排序，从小到大选取，直至所有机组出力的加和等于预测值时停止。中间还必须考虑各机组爬坡速率对其在某一时段出力上（下）限的影响。

然而，由于电网的各条线路上的负荷有一个上限，一旦超过，就会出现输电阻塞的情况。出现阻塞时，先在不超过线路潮流值上限的前提下调整出力分配方案，使电阻塞消除。如果做不到，还可以使用安全裕度输电，即在遇到紧急情况时，潮流的绝对值可以超过限值的一定百分比，但要使每条线路上的潮流的绝对值超过限值的百分比尽可能的小。如果以上措施均不能解决阻塞问题，则被迫采取在用电侧拉闸限电措施，但会造成人民财产的巨大损失。

在出现改变出力分配的预案时，对一些通过竞价取得发电权的发电

容量(序内容量)不能出力的和一些在竞价中未取得发电权的发电容量(序外容量)要在低于对应报价的清算价上出力的情况,予以适当的经济补偿,但要注意公平原则。以上的经济补偿称为阻塞费用。网方当然希望尽量减少阻塞费用。

题中给出了某电网8台发电机组,6条主要线路的一些相关数据。要求首先根据各机组分力方案和各线路潮流值的关系,找出两者间的函数关系,再设计出一种公平合理的阻塞费用计算规则。在此基础上,对两组具体数据(负荷需求982.4MW和1052.8MW)给出出力分配方案并根据安全原则讨论之。

二、基本假设

(1)电力传输过程中,在线路中的功耗(即网损)以及电网中的无功功率损耗均不计。

(2)时段间机组的输出功率改变时,总是以最大爬(下)坡速率改变到需要的输出功率,之后在AGC的辅助服务下保持稳定。

(3)6条线路是相对独立,彼此之间影响不大。

(4)题中所给的数据都存在着一定的随机误差。

三、变量及符号声明

x:机组的出力;

y:线路的潮流值;

i:输电线路,$i=1,2,\cdots,6$;

j:发电机组,$j=1,2,\cdots,8$;

p:清算价;

$p(x_j)$:x_j 所在段的段价;

a_j:分配预案中发电机组 j 的出力;

S:阻塞费用;

l_i:线路 i 的潮流限值;

r_i:线路 i 的相对安全裕度;

λ_j:阻塞贡献系数;



D:预测的需求;

其余一些不常用的变量在文章中用到时会在后面注明。

四、模型的建立、求解和分析

1. 对有功潮流和机组出力关系的初步分析

(1)拟合曲线猜想

我们发现机组的出力和各线路的潮流值的量纲均为 MW,于是首先通过量纲分析法,猜想两者之间应该是一个线性的关系,于是先对单变量情况进行验证,对于某个 x_j,改变其数值,在其他分量不变的前提下,观察 y_i 的值如何对应的发生变化,绘制出 y_i 对 x_j 的散点图,并拟合曲线。

经过拟合,共得到 48 张图,我们发现的多数的单变量(机组的出力)对线路潮流值的影响都很好的符合线性关系。由于篇幅限制,仅任选两张较典型的附在此处,以说明问题。

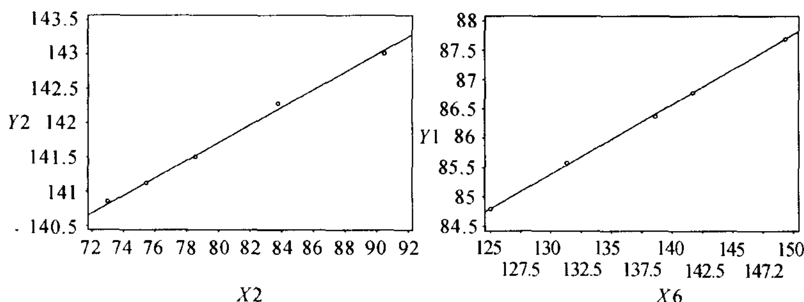


图 1

(2)多元线性回归模型的建立和求解

验证了线性关系后,在假设 3 的前提下,考虑从 R^8 到 R^6 的一个多元映射,其中像空间的每个方向上对每个分量都是线性的函数。由于技术和计算成本的限制,且出于近似要求,我们可以认为表 2 中所给数据与理想值间存在随机误差,同假设 4,记为 ε,假定 $\varepsilon \in N(0, \delta^2)$。根据提出的线性猜想,我们建立如下的数学模型:

$$Y = XB + \varepsilon$$

其中，

$Y=[y_1,\cdots,y_6]^T$，$y_i(i=1,\cdots,6)$　表示第 i 条线路上的潮流值；

$X=[1,x_1,\cdots,x_8]^T$，$x_j(j=1,\cdots,8)$　表示第 j 个机组的出力。

$$\hat{B}=\begin{pmatrix}\hat{b}_{1.0} & \cdots & \hat{b}_{6.0}\\ \vdots & \ddots & \vdots\\ \vdots & & \vdots\\ \hat{b}_{1.8} & \cdots & \hat{b}_{6.8}\end{pmatrix}$$

$\hat{b}_{i,j}(i=1,\cdots,6;j=1,\cdots,8)$表示第 j 个机组的出力对第 i 条线路上潮流值的边际影响。

$\varepsilon=[\varepsilon_1,\cdots,\varepsilon_6]^T$，$E(\varepsilon)=0$

对于所列模型，对题中所给数据进行多元线性回归分析。

由题中表 1，表 2 所给的 33 组数据，引入矩阵

$$Y=\begin{pmatrix}y_{1.0} & \cdots & y_{6.0}\\ \vdots & \ddots & \vdots\\ \vdots & & \vdots\\ y_{1.32} & \cdots & y_{6.32}\end{pmatrix}\qquad X=\begin{pmatrix}1 & x_{0.1}\cdots & a_{0.8}\\ \vdots & \cdots & \vdots\\ \vdots & \cdots & \vdots\\ 1 & x_{32.1}\cdots & x_{32.8}\end{pmatrix}$$

用最大似然估计法估计参数$\hat{B}$，可得：

$$\hat{B}=(X'X)^{-1}X'Y$$

$$=\left(\begin{pmatrix}1 & 120\cdots & 90\\ \vdots & \cdots & \vdots\\ \vdots & \cdots & \vdots\\ 1 & 120\cdots & 120.44\end{pmatrix}^T\begin{pmatrix}1 & 120\cdots & 90\\ \vdots & \cdots & \vdots\\ \vdots & \cdots & \vdots\\ 1 & 120\cdots & 120.44\end{pmatrix}\right)^{-1}\cdot$$

$$\begin{pmatrix}1 & 120\cdots & 90\\ \vdots & \cdots & \vdots\\ \vdots & \cdots & \vdots\\ 1 & 120\cdots & 120.44\end{pmatrix}^T\begin{pmatrix}y_{1.0} & \cdots & y_{6.0}\\ \vdots & \ddots & \vdots\\ \vdots & & \vdots\\ y_{1.32} & \cdots & y_{6.32}\end{pmatrix}$$

经计算得到的函数如下：

$$y_1=0.0826x_1+0.0478x_2+0.0528x_3+0.1199x_4-0.0257x_5+0.1216x_6+0.1220x_7-0.0015x_8+110.48$$

$$y_2 = -0.0547x_1 + 0.1275x_2 - 0.0001x_3 + 0.0332x_4 + 0.0867x_5 - 0.1127x_6 - 0.0186x_7 + 0.0985x_8 + 131.35$$

$$y_3 = -0.0694x_1 + 0.0620x_2 - 0.1565x_3 - 0.0099x_4 + 0.1247x_5 + 0.0024x_6 - 0.0028x_7 - 0.2012x_8 - 108.99$$

$$y_4 = -0.0346x_1 - 0.1028x_2 + 0.2050x_3 - 0.0209x_4 - 0.0120x_5 + 0.0057x_6 + 0.1452x_7 - 0.0763x_8 + 77.612$$

$$y_5 = 0.0003x_1 + 0.2428x_2 - 0.0674x_3 - 0.0412x_4 - 0.0655x_5 + 0.0700x_6 - 0.0039x_7 - 0.0092x_8 + 133.13$$

$$y_6 = 0.02376x_1 - 0.0607x_2 - 0.0781x_3 + 0.0929x_4 + 0.0466x_5 - 0.0003x_6 + 0.1644x_7 + 0.0004x_8 + 120.85$$

(3)对变量的显著性检验和模型的进一步改进

对每个因变量 x_j 与 y_i 的关系进行显著性分析。

对于存在 m 个自变量,由已知的 n 组实验数据进行的多元线性回归分析,可根据以下理论[1],对其显著性进行假设检验:

在显著性水平 $\alpha = 1\%$ 下检验 $H_{ij}: b_{ij} = 0, j = 1, \cdots, m$

若 H_{ij} 为真,

$$t_{ij} = \frac{\hat{b}_{ij}}{\sqrt{c_{jj}} \cdot s} \sim t(n - m - 1), j = 1, 2, \cdots, m$$

而当 H_{ij} 不成立时,t_{ij} 有变大的趋势,因而应该取双侧拒绝域,故当 $t_{ij} > t_{\frac{\alpha}{2}}(n - m - 1)$ 时,否定 H_{ij},则从回归方程中删除那些不起作用的自变量。

借助 SAS 检验分析,按以上标准删除了一些变量,并重新对此进行多元线性回归分析,得到:

$$y_1 = 0.0829x_1 + 0.0484x_2 + 0.0530x_3 + 0.1200x_4 - 0.0254x_5 + 0.1221x_6 + 0.1215x_7 + 110.15$$

$$y_2 = -0.0547x_1 + 0.1276x_2 + 0.0332x_4 + 0.0867x_5 - 0.1126x_6 - 0.0187x_7 + 0.0986x_8 + 131.30$$

$$y_3 = -0.0700x_1 + 0.0607x_2 - 0.1570x_3 - 0.0101x_4 + 0.1240x_5 - 0.2020x_8 - 108.48$$

$$y_4 = -0.0346x_1 - 0.1028x_2 + 0.2050x_3 - 0.0209x_4 - 0.0120x_5 + 0.0057x_6 + 0.1452x_7 - 0.0763x_8 + 77.612$$

$$y_5 = 0.2427x_2 - 0.0648x_3 - 0.0412x_4 - 0.0655x_5 + 0.0699x_6 - 0.0038x_7 - 0.0093x_8 + 133.22$$

$$y_6 = 0.2375x_1 - 0.0607x_2 - 0.0781x_3 + 0.0929x_4 + 0.0466x_5 + 0.1664x_7 + 120.86$$

再进行一次显著性假设检验，此时，对所有的 x_{ij}，除 x_7 对 y_5 的影响外，其余的显著性均能达到 0.0001 的显著水平，但所有的显著性均符合 $\alpha = 1\%$ 的显著性水平。

比较两次的结果，在删除那些不显著变量的前后，其余变量的系数（边际影响）和常数项的变化均在千分之一以下，这又对以上理论给出了很好的验证。

(4)更深的思考

我们对解析式中常数项的物理意义一直耿耿于怀。按照这个计算，当所有的发电机组都不工作时，按这个公式，每条输电线路上仍将有相当大的功率传输。

我们试图寻找出题中电网的拓扑结构，希望能经过此寻找到答案，但是失败了。又查了很多资料仍得不到好的解释，因为实际电网中的有功潮流要根据电网的实际情况，结合基尔霍夫定律和平衡性的考虑，进行大量的分析和计算才能得到，而题中给的显然是经过简化后的模型，所以我们有理由认为题中要求的近似表达式只是一个近似的经验公式，只在电网正常工作的范围内能比较好的近似描述各线路上的有功潮流和各个发电机组出力的关系。因此这里的常数项的出现是合理的，并非我们凭空臆造的。

2.简明合理且公平的阻塞费用计算法则

按照题意，每个机组的报价是一个与出力有关的阶梯函数，其中每一个阶梯就是一个段位，该阶梯的纵坐标即段价（从负到正单调增加），而横坐标方向上的长度表示段容量，如图 2 所示。

对于重新分析机组出力的公平补偿问题，通过查阅资料我们了解到，



在我国、英国、美国和澳大利亚等国家均采取了不同的阻塞管理模式，对于阻塞费用的分摊大体上有平均分摊法和使用程度分摊法[2]：

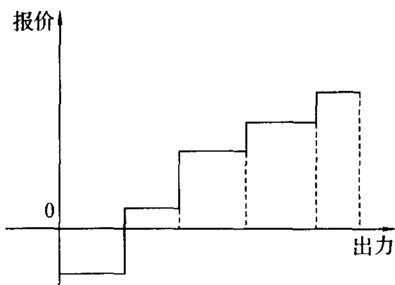


图 2

平均分摊法即俗称的邮票法，采用按在新分配下改变的发电力功率大小分摊给各发电商，不考虑市场成员对网络约束的影响，该方法的主要优点是简单。而最大的缺点则是不能公平、合理的分摊阻塞费用，难以体现题中要求的公平性，本文在此不予考虑。

而使用程度分摊又包括影子价格法，边际微增分配法，灵敏度分析法，支付因子法，Lagrange 松弛算法等多种方法，这里我们可以证明，完全满足全部这些衡量标准的完全公平程度分摊是不存在的。

定理 1　综合满足 n 种程度分摊评判方法对 $m(m\geqslant3)$ 个机组进行损失综合分摊评定，不可能实现完全的公平。

证明：完全的公平原则必须满足以下几条，不妨以公理的形式给出：

公理 1　各种分摊对 m 个机组的程度分摊的排预订可以实现。

公理 2　定义 $(x\geqslant y)_i$ 表示用第 i 种评判方法 x 的责任分摊大于 y。则如果 $\forall i$，都有 $(x\geqslant y)_i$，于是必有 $(x=y)\Leftrightarrow\forall i,(x=y)_i$，即尊重评判方法的一致分摊排序。

公理 3　一种评判方法在两次评判中不会得到不同的分摊排序。

公理 4　如果两次评判中，每种评判方法对 x 和 y 的评分摊排序相同，则完全公平的原则下 x,y 的排序不改变。

公理 5　不存在垄断评判法，即不存在 I，使得 $(x\geqslant y)_i$ 必然决定 $x\geqslant y$。

不难证明，这 5 条公理的体系不相容。可以从前面 4 条推出存在一种垄断的判别方法，即与第 5 条公理矛盾。具体可以仿照著名的 Arrow 不可能定理[3]证明的方法，此处限于篇幅，不再赘述。

基于上述定理，我们知道想做到完全公平的分摊是不可能的，所以通



过比较筛选得到几个比较行之有效而且便于理解的分摊方法：

(1)Pool 模式的标准法[2]：

记分配预案中第 j 个机组的出力为 a_j，调整后的出力为 x_j，x_j 所在段对应的段价为 $p(x_j)$，$p(x_j)$为图 1 所示的分段函数，清算价为 p。

补偿原则

对于序内不能出力部分，以清算价高于报价的金额给予补偿

$$(p - p(x_j))\cdot(a_j - x_j)$$

对于序外容量调整后出力部分，以报价高于清算价的金额给予补尝

$$(p(x_j) - p)\cdot(x_j - a_j)$$

则阻塞费用的表达式为：$S = \sum_{j=1}^{8}(p(x_j) - p)\cdot(x_j - a_j)$

(2)改进的 Pool 方法

方法(1)的分摊方法在一定程度上体现了对待序内容量不能出力和报价高于清算价的序外容量出力的公平性，但是没有很好的体现阻塞程度分摊的公平思想，在此我们提出两种改进的模式(后文对题中所给数据的计算可以验证两种模式得到的结果差别很小)：

①基于 Pool 模式的阻塞量分摊法

对于机组 j 的阻塞贡献λ'_j 进行计算，我们先考虑线路 i 的阻塞量占总阻塞的比重 ω_i，

则 $$\omega_j = \frac{\Delta y_i}{\sum_i \Delta y_i},$$

其中$\Delta y_i = \max\{0, y_i - l_i\}$表示此线路上的阻塞值，其中 l_i 表示此线路的潮流限值，y_i 为有功潮流；

然后计算这条线路上机组 j 有效有功潮流占此线路所有有效有功潮流的比例σ_j，$\sigma_j = \dfrac{\eta_j}{\sum_{j=1}^{8}\eta_j}$，$\eta_i = \max\{b_{i,j}\cdot x_j, 0\}$ 表示有效阻塞(题中给出的数据都是超出 l_i 的，故可以如此表示；若有低于 $-l_i$ 的可以另外表示)。

则 $$\lambda'_j = \sum_{i=1}^{6}\omega_i\cdot\sigma_j = \sum_{i=1}^{6}\frac{\Delta y_i}{\sum_{i=1}^{6}\Delta y_i}\cdot\frac{\eta_j}{\sum_{j=1}^{8}\eta_j},$$

对 $\lambda'_j(j=1,\cdots,8)$ 进行归一化处理得到新的有效阻塞：

$$\lambda_j = \frac{\lambda'_j}{\sum_{j=1}^{8} \lambda'_j},$$

显然，$\sum_{j=1}^{8} \lambda_j = 1$

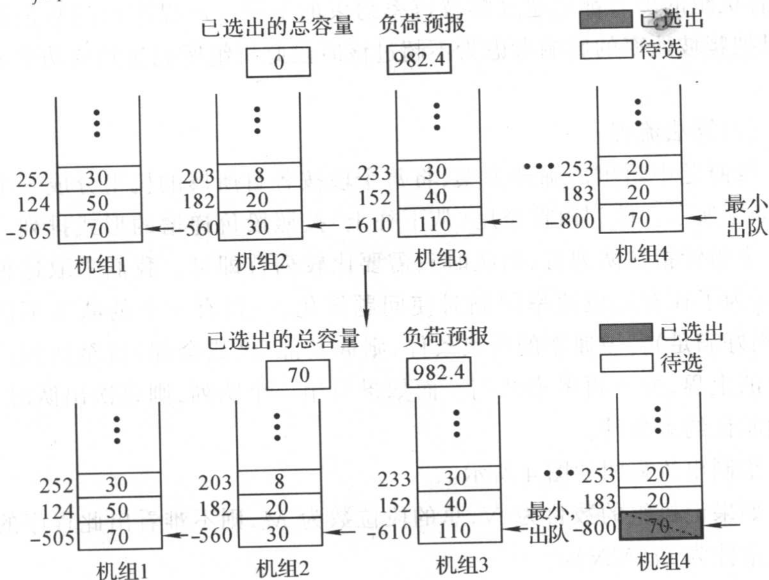


图 3

因为各机组对阻塞都有贡献，所以造成阻塞的机组也应该付出一定的代价，故我们认为不需把 Pool 模式标准算法所计算的补偿全部支付，只需支付$(1-\lambda_j)$的比例即可。改进后的阻塞费用为：

$$S' = \sum_{j=1}^{8} (1-\lambda_i)(p(x_j)-p)\cdot(x_j-a_j)$$

②基于 Pool 模式的影子价格法

理论基本同上，只是线路 i 上机组 j 有效阻塞占此线路阻塞值的比例 σ_j 的计算略有不同，这里使用的是机组 j 的有效边际贡献阻塞（即影子价格），而不是有效贡献阻塞量。记影子价格法得到的阻塞费用为 S''。

3. 一个时段机组出力的分配预案

(1)爬坡速率的约束

我们知道爬坡速率决定了各机组在下一时段的出力范围。对题中所给的两个负荷,我们先不考虑爬坡速率的约束。按照题中给出的选取标准,价格最低的先选,即用贪婪法进行计算。通过对结果的观察我们发现,各机组的出力都远超过爬坡速率约束的下限。所以本题的数据我们可以把爬坡速率的影响考虑为不超过该时段此机组所能达到的功率的上限。

(2)算法流程

按照题中给出的筛选方案,将每个段按各自所属的机组分成 8 个队列,如图所示。队列由首至尾(从下往上)按照段价递增的形式排列。各用一个指针指在队列首,出队时只需要比较列首即可。我们采取这种方法,是为了在有爬坡速率限制时使问题简化,一旦有一个的队头不满足(或刚好满足)爬坡速率的约束条件,则取其部分(或全部)直至达到约束条件的上界,并不再考虑队列。而如果只用一个队列,则每次出队时,都必须验证约束条件。

绘制出流程图如图 4 所示。

如果记机组的数目为 N,总的段位数为 M,则不难看出此程序的计算复杂性为 $O(MN)$。

(3)定理 2 各机组时段报价已定时,按照惯例,结算价应该按照所取的所有段的价格的最大值计算。此时,题中所给的方法,即按段价从低到高选取各机组的段容量或其部分,直到它们之和等于预报的负荷,并以最后一个选入的段价作为清算价,可以保证总费用最低。

证明(反证法):记题中给的方法为方法 1,假设存在方法 2 得到的满足负荷的分配的总费用低于方法 1,由于总量不变(即为预测负荷),则方法 2 的结算价小于方法 1 的清算价 P,那么方法 2 中所取段的价格都小于 P,根据方法 1 的规则可知,方法 2 所取的分配中的所有段都包含于方法 1 的分配中,故方法 2 的分配必然小于预测负荷减去方法 1 中段价为 P 的容量,显然不可能满足预测负荷,所以假设不成立,得证。



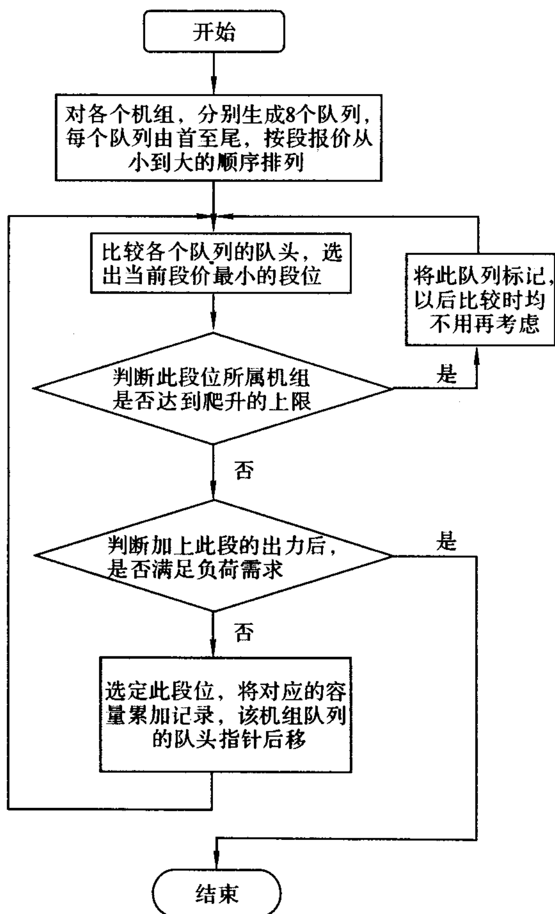


图 4

4.考虑潮流限值的出力分配方案

在上一问题的基础上，计算是否出现输电阻塞(即存在线路有功潮流超过限值)：

①如果不阻塞，则无需付阻塞费用，且按照定理 2，此时有总费用最低。

②如果出现有阻塞的情况，则首先判断是否可以通过调整使阻塞消

除,即判断在满足给定负荷要求下,是否可以使线路上有功潮流都不超过限值

这里,我们利用函数 $\min\limits_{x_1,\cdots,x_8}\left(\max\limits_{1\leqslant j\leqslant 6}(|y_i(x_1,\cdots,x_8)|-l_i)\right)$进行判断。

所谓 min,max 函数,指的是所有函数的最大值在自变量所有可行取值范围内的最小值。

求解规划 1:

$$\min_{x_1,\cdots,x_8}\left(\max_{1\leqslant j\leqslant 6}(|y_j(x_1,\cdots,x_8)|-l_j)\right) \qquad \text{(规划 1)}$$

$$\text{s.t}\quad \inf(x_j)\leqslant x_j\leqslant \sup(x_j),\quad j=1,2,\cdots,8 \qquad \text{(爬坡速率的约束)}$$

$$-l_i\cdot(1+r_1)\leqslant y_i\leqslant l_i\cdot(1+r_i),\quad i=1,2,\cdots,6$$

(安全裕度上限的约束)

$$\sum_{j=1}^{8}\sum_{n=1}^{10}x_{nj}=D, \qquad \text{(总的出力与预测的需求相等)}$$

$y_j(x_1,\cdots,x_8)$表示此时第 j 条线路上的有功潮流值,$\inf(x_j)$、$\sup(x_j)$分别表示根据上一段的出力值,结合爬坡速率,得到的 x_j 的上下限,(以下相同)。

③如果得到结果小于等于零,即每条线路的有功潮流均未超过上限,则在线路潮流的限值条件下进行调整可以满足题中要求,执行规划 2,求出在公平对待下调整所需阻塞费用的最小值。

$$\min S \qquad (S\text{ 为所求的阻塞费用}) \qquad \text{(规划 2)}$$

$$\text{s.t.}\quad \inf(x_j)\leqslant x_j\leqslant \sup(x_j)\quad j=1,2,\cdots,8 \qquad \text{(爬坡速率的约束)}$$

$$-l_i\leqslant y_i\leqslant l_i \qquad i=1,2,\cdots,6 \qquad \text{(有功潮流限值的约束)}$$

$$\sum_{j=1}^{8}\sum_{n=1}^{10}x_{nj}=D \qquad \text{(总的出力与预测的需求相等)}$$

这个规划的目标函数在前面曾给出三种不同的表达。而且目标函数中含有分段函数,我们分别利用下列方法对其进行求解:

● 使用 lingo8.0 完整版,将分段函数在程度中利用 if 语句表示并求解,速度比较慢,大概要 20 秒左右。

● 对于分段函数的表示,加入若干个控制变量,将目标函数改写成一个 0—1 规划并求解,时间比(1)要快。

● 利用分枝定界的思想将规划 2 转化成若干个线性规划进行求解,



这个方法速度比较快,不过人工处理的工作量比较大。

由于题中给出的条件有一定的局限性,我们认为在现实生活中网络线路很复杂的情况下,可以考虑用遗传算法去求解,因为网方要尽快计算出结果(至少在 15 分钟内完成)。

④若得到的结果大于零,则在线路潮流的限值条件下不能满足负荷需求,必须使用线路的安全裕电输电。

此时为了保证安全第一的原则,题中要求使每条线路上潮流超限值的百分比尽量小,我们认为在超限百分比相等的情况下,安全裕度较大的线路危险性较小,比如对于线路 2 和线路 3,安全裕度分别是 18% 和 9%,那么如果都是超过上限 10MW,对于两者来说是大不一样的。因此,应该考虑各线路安全裕度上限的影响,我们提出以下两种模型:

(1)多目标规划:

$$\min\left\{\max\left\{\frac{|y_i|-l_i}{l_i},0\right\}\right\} \qquad i=1,2,\cdots,6 \qquad \text{(规划 3)}$$

$$\text{s.t.}\quad \inf(x_j)\leqslant x_j\leqslant \sup(x_j) \qquad j=1,2,\cdots,8 \qquad \text{(爬坡速率的约束)}$$

$$-(1+r_i)l_i\leqslant y_i\leqslant(1+r_i)l_i \qquad i=1,2,\cdots,6$$

(有功潮流安全裕度限值的约束)

$$\sum_{j=1}^{8}\sum_{n=1}^{10}x_{nj} = D \qquad \text{(总的出力与预测的需求相等)}$$

目标函数中,MAX 表示如果超过上限,则取超过部分的值,否则取 0。对于每个目标的权值我们用 $\dfrac{\frac{1}{r_i}}{\sum_{i=1}^{6}\frac{1}{r_i}}$ 来表示,将规划 3 转化为一个单目标规划规划:

$$\min\sum_{i=1}^{6}\left(\frac{\frac{1}{r_i}}{\sum_{i=1}^{6}\frac{1}{r_i}}\max\left\{\frac{|y_i|-l_i}{l_i},0\right\}\right) \qquad \text{(规划 4)}$$

$$\text{s.t.}\quad \inf(x_j)\leqslant x_j\leqslant \sup(x_j) \qquad j=1,2,\cdots,8 \qquad \text{(爬坡速率的约束)}$$

$$-(1+r_i)l_i\leqslant y_i\leqslant(1+r_i)l_i \qquad i=1,2,\cdots,6$$

(为有功潮流安全裕度限值的约束)



$$\sum_{j=1}^{8}\sum_{n=1}^{10}x_{nj}=D \qquad \text{(总的出力与预测的需求相等)}$$

(2) 用函数 $\min(\max\limits_{1\leqslant i\leqslant 6}\frac{|y_i|-l_i}{l_i\cdot r_i})$ 来衡量

$$\min(\max_{1\leqslant i\leqslant 6}\frac{|y_i|-l_i}{l_i\cdot r_i}) \qquad \text{(规划 5)}$$

s.t. x_j 的下限 $\leqslant x_j \leqslant x_j$ 的上限　$j=1,2,\cdots,8$　(爬坡速率的约束)

$-l_i(1+r_i)\leqslant y_i\leqslant l_i(1+r_i)$　$i=1,2,\cdots,6$

(安全裕度上的约束)

$$\sum_{j=1}^{8}\sum_{n=1}^{10}x_{nj}=D \qquad \text{(总的出力与预测的需求相等)}$$

这种函数的衡量意义和优越性在后文具体的数据计算中做了分析。

如果在此约束条件下仍然找不到可行解,则表示必须采取拉闸限电措施以确保线路的安全,但是本题中不会发生拉闸的现象。

五、对题中所给数据的求解

(1)预报的下一时段负荷需求为 982.4MW 时,按照前面的程序求解,得分配预案如表 1 所示:

表 1　预报负荷需求为 982.4MW 的出力分配预案

机组编号	1	2	3	4	5	6	7	8
出力(MW)	150.0	79.0	180.0	99.5	125.0	140.0	95.0	113.9

此时的清算价为 303 元/MWh。

各条线路上的潮流值如表 2 所示。

表 2　按表 1 的分配预案时各条线路上的潮流

线路编号	1	2	3	4	5	6
潮流值(MW)	173.4	141.0	−151.0	103.5	136.8	168.5

比较此时的分配预案(表 1)和各线路的潮流限值(附表 6),发现线路



5、6上均有超过限值的情况。

执行规划1，得到值小于0，表示此时可以通过线路上的调节，消除超过有功潮流上限传输的现象，无需使用裕度发电。

求阻塞费用最小，分别利用Pool模式标准法(S)，基于Pool模式的阻塞分摊法(S')及影子价格法(S'')，执行规划2，得到表3。

表3 预报为982.4MW时的分配及补偿方案

机组编号		1	2	3	4	5	6	7	8	最小阻塞费用(元)
改进后的出力(MW)		151.7	88.0	228.0	80.0	152.0	95.7	70.0	117.0	
15分钟获得的补偿(元)	Pool模式标准法	77.0	432.0	636.0	234.0	1397.3	1593.3	325.0	0.0	4694.5
	阻塞分摊法	54.6	377.4	592.9	201.8	1350.9	1319.1	269.1	0.0	4165.8
	影子价格法	59.1	361.4	606.7	194.9	1353.2	1368.6	254.1	0.0	4198.1

观察数据可以发现，在给出的三种判断标准下，重新分配后，各个机组的出力完全相同，最小阻塞费用的差别只是由补偿的比例不同而造成的。另外我们又根据Lagrange松弛算法和灵敏度分析法对题给数据进行了具体计算，发现不同的计算方案求得的一个时间段内阻塞费用与改进后的两种方法(S′,S″)差别很小(<100元)。这很好地说明了虽然不能做到完全公平，但在实际生产中利用这些方法还是完全合理和可行的。

(2)预报的下一时段负荷需求为1052.8MW时，同样按照前面的程序求解，得分配预案如表4所示。

表4 预报负荷需求为1052.8MW时的出力分配预案

机组编号	1	2	3	4	5	6	7	8
出力(MW)	150.0	81.0	218.2	99.5	135.0	150.0	102.1	117.0

此时的清算价为:356元/MWh。

各线路上的潮流如表5所示:

表5 按表一的分配预案时各条线路上的潮流

线路编号	1	2	3	4	5	6
潮流值(MW)	177.3	141.2	-156.2	111.9	134.8	167.1

此时仍是线路 1,5,6 不满足限制条件。执行规划 1,结果大于 0,表示需要尝试在裕度条件下输电。

分别执行规划 4 和规划 5,比较两者的结果。

表 6　规划 4 下的各机组的分析

机组编号	1	2	3	4	5	6	7	8
出力(MW)	134.0	75.0	228.0	99.5	152.0	92.3	117.0	
15 分钟获得的补偿(元)	416.2	83.3	0.0	0.0	654.5	30.0	236.2	0.0

此时,只有线路 1 上超过了安全上限 10.2MW,占 6.17%。

根据 Pool 模式标准法,算出阻塞费用 1420 元。

表 7　规划 5 下的各机组的分配

机组编号	1	2	3	4	5	6	7	8
出力(MW)	153.0	88.0	228.0	99.5	152.0	113.2	102.1	117.0
15 分钟获得的补偿(元)	99.8	243.3	0.0	0.0	654.5	1683.6	0.0	0.0

此时,线路 1,5,6 均超过了安全上限,线路 1 超过了 8.5MW,占 5.14%;线路 5 超过了 0.2MW,占 0.14%;线路 6 超过了 5.4MW,占 3.3%。

根据 Pool 模式标准法,算出阻塞费用 2681.1 元。

比较表 6 和表 7 的结果:

对于线路上的危险,显然,边际危险是随着超额的百分比递增的,而一个电网的安全运行,取决于每条线路的安全性,即若有一条线路出现危险,则电网将不能正常运行。那么,我们有理由认为,线路 1 上超限 6.17%远比 5.14%危险,而规划 5 利用了 MIN(MAX)函数,将线路上的危险分散了,降低了整个网络的危险性。由此可见,规划 5 的改进是有效的。



六、模型分析

1. 对拉闸限电的分析

我们可以求解在裕度允许范围内,该电网的最大供电量,建立规划:

$$\max \sum_{j=1}^{8} x_j$$

$$\text{s.t} \quad \inf(x_j) \leqslant x_j \leqslant \sup(x_j) \quad j=1,2,\cdots,8 \quad (\text{爬坡速率的约束})$$

$$-l_i \cdot (1+r_i) \leqslant y_i \leqslant l_i \cdot (1+r_i) \quad i=1,2,\cdots,6$$

(安全裕度上限的约束)

得到的结果为 1094.6MW。但通过分析可知该数值不是受到线路潮流限度的限制,而是因为机组的爬坡速率不够大。

如果我们考虑爬坡速率足够大的情况,同等条件下求出的最大供电量为 2771MW,可见要解决供电紧缺的情况,提高爬坡速率是关键步骤之一。

2. 灵敏度分析

我们针对上述规划做灵敏度分析,并通过 Matlab 做出灵敏度分析图像做比较分析,可以发现,线路 2～6 的灵敏度分析图像均为一条水平的线,只有线路 1 是整个网络的"瓶颈",它的灵敏度图像见图 5。

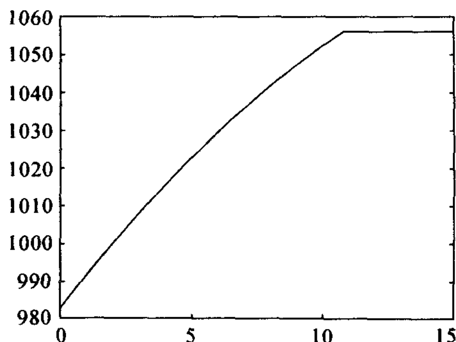


图 5

横坐标为线路 1 上限的增加量,纵坐标为相应的最大供电量的变化。显然,在开始相当长的一段内,线路 1 都是瓶颈。想提高题中给出的电网的供电能力,对线路 1 的投资扩建或其它各种技术性改建都是很必要的。

我们比较了各机组爬坡速率的灵敏度,可以知道机组 5,8 的灵敏度

比较大,若想提高整个网络的供电能力,对它们的改良值得优先考虑。

3.算法稳定性分析

我们考虑求最小阻塞费用的规划(只考虑 Pool 模式标准化),因为题中给的数据都不是准确值,所以我们用 Monte Carlo 算法的思想,对题给数据中随机性比较大的数值在其 ±1% 的范围内去满足正态分布的随机扰动,我们利用计算机给出 1000 个随机生成的数据进行仿真模拟,其中无解的比例小于 15%,统计量 $\sqrt{\sum_H\left(\frac{Z_H - Z_{OPT}}{Z_{OPT}}\right)^2}$ 的值小于 10%。其中 Z_H 为仿真中得到的最优值,Z_{OPT} 为论文中得到的最优解。

论文中我们设计的其他模型都是线性规划或者可以通过分枝定界的思想转换为若干个线性规划,很易求解。所以整体来说,我们设计的算法比较稳定,可以接受。

七、模型的评价

1.优点评述

(1)使用显著性假设检验使表达式更准确。

(2)MIN(MAX)函数的使用,使在讨论安全裕度输电问题时更加合理。

(3)对于阻塞费用的计算我们提出不同的模型并做了对比和评述,使公平性得到了更好的体现。

(4)在解决题目所问的问题时,可以利用现成的模型,很方便的找到电网的瓶颈,为以后的改造提供参考依据。

(5)稳定性检验验证了模型的稳定性,为算法的正确性提供了有力的依据。

2.模型改进

(1)考虑机组爬坡阶段的功率的改变

机组不可能在瞬间从一个输出功率跳到第二个,在每个段位之间改变时,需要经过一个爬坡的阶段,使得总的输出功比按照理论计算的要少一些。

我们计算赔偿费用,即阻塞费用时提出的方法没有考虑爬坡速率的变化过程,因为本来的赔偿方式就是一种估算的方法。考虑实际情况应该是15分钟内发电功率的变化对时间的一个积分,这段时间的功才是发电厂实际的做功。所以若想得到更准确的分析,可以考虑把爬坡用的时间考虑进去。

(2)这里,题中已经给出了下个时段的负荷预测值。实际情况中,网方需要根据用户的实际需要,对下个时段的负荷进行预测。这里可以根据历史数据和实际情况的分析,给出负荷预测函数。如果能给出一个较好的预测函数,网方在选择电价上将占有更大的主动,也有利于对整个供电系统进行宏观调控。

(3)题中将将发电方作了简化,其实在实际生产中,火电站、水电站、核电站等之间还是有着很显著的差异。如果遇到实际问题时,需要综合考虑各种发电厂的特点,以便更好的利用时段,价格等方面的因素。

(4)我们查阅过一些实际数据(比如华东电网的组成),比题中的模型要复杂得多,所以这里适用的方法在给出条件复杂到一定程度时便不能很好的解决问题,这时就要考虑一些启发式算法去近似求解,如遗传算法等。

八、引用文献

[1]贾敬、桂占吉,数学建模与数学实验,1998年7月
[2]杨洪明、段献忠、何仰赞,阻塞费用的计算和分摊,电力自动化设备,2002年5月,22卷第5期
[3]高宏业,西方经济学(微观部分),北京:人民大学出版社,2001

点评

随着近年来全国一片拉闸限电之声以及"厂网分开,竞价上网"等电力体制改革措施的推出,引发了人们对电力行业的关注和利用数学模型调节电力市场的期盼。本题即是在上述背景下产生的。对大部分选手而言,问题中的许多术语是陌生的,能否尽快吃透相关的基本概念,进入问

题情景，对整个比赛进程影响甚大。

本题的求解大致包括以下三个方面。

(1)得到有功潮流和机组出力之间的关系，这是一个统计分析问题，对给出的33组数据用多元线性回归可得满意的结论。值得思考的是回归方程中常数项的选择上，尽管物理意义似乎不支持该项假设，但考虑到实验电网的局部性，试验数据的随机性以及统计检验结论，保留常数项不乏可取之处。

(2)求出出力分配预案，这是一个算法设计问题，根据交易规则和爬坡速率的限制，不难写出相应的算法。预案由报价惟一确定，尚不存在调整优化问题。

(3)出现阻塞时的调整方案，这是一个非线性规划的优化计算问题。注意到按题意，对“输电阻塞”、“安全裕度供电”和“拉闸限电”三种模式下优化的对象是不同的，对应于规划的目标函数是各不相同的。其中潮流超过限值时要求各条线路超限的百分比最大值极小化的思想，以及如何在目标中正确体现安全和经济两者关系尤为重要。

这篇论文对以上三个方面的求解都是比较精彩的，另外本文还对如何分摊阻塞费用进行了详细的讨论，给出了多种规则，并进行了比较分析，这是它的特色之处。

本文获2004年全国大学生数学建模竞赛全国一等奖。

第二部分

美国大学生数学建模竞赛

Analysis of asteroid impact

杨　骏　刘自强　张子健

Summary

We develop a method to assess the possible consequences of an asteroid striking in Antarctica. Three models are included: Nuclear Simulation Model, Ecosystem Model and Sea Level Model.

Nuclear Simulation Model is constructed to assess the direct damage of the asteroid on the basis of nuclear tests, because no other natural disasters can be compared with the impact except nuclear wars. We also modify the initial model to fit for the special landform of Antarctica. Using this model, we can estimate the size of the crater created by any asteroid, given its diameter and velocity.

We concentrate on the long-term effect of the impact instead of short-term damage it produces. The problems of food production and sea level change by melting ice sheet are highlighted.

Food production variance due to the decline of solar energy is estimated in Ecosystem Model. The famous Volterra's predator-prey model is included to describe the changes of two typical species in ocean: algae and shrimp. Our computer simulation shows that great impact leads to destruction of ecosystem, while small one causes a decline and 'oscillation' of food production. Possible human casualties are discussed according to the distribution of food supply.

Sea Level Model is applied to estimate sea level change. Multiple factors, such as greenhouse gases, vaporization of seawater, melting of ice sheet, are included in



our model. As a result of simulation, we learn the main cause of melting ice is not the direct impact. Instead, the 'nuclear winter' and the enhanced greenhouse effect contribute a lot to the unexpected variance of sea level.

Throughout our modeling, four asteroids A, B, C and D, varying in diameter and velocity, are selected as our samples in order to learn the different effect of asteroids of different properties. An additional sample B* will enable us to learn the different consequences produced by the impact in place other than Antarctica.

The impact of a 'threshold asteroid', with 1 km in diameter and 20km/s in velocity, is estimated. The impact greater than it is likely to cause global catastrophe, while the others below it only causes local damage. Finally, parameter sentivities are tested.

Analysis of Possible Hazard due to Asteroids

It is believed that, the Earth has been continuously impacted by countless asteroids, comets and meteoroids during the long history of its formation. This can be easily proved by the records of impacts in the Earth's geological history. The 1908 Tunguska explosion caused by atmospheric disruption of a stony asteroid and the Meteor Crater in Arizona are the typical events among them. Moreover, scientists believe that, the extinction of dinosaur was due to the impact of an asteroid with diameter of 10km (so called K/T impact). In fact, this impact which happened 65 million years ago wiped out half the living species on earth, resulting in the destruction of ecosystem.

Especially, the Near Earth Objects (NEOs) on the order of 1000m in diameter draw more attention, since they are believed to be capable of producing global disasters. Their civilization threatening consequences compensate well for their low frequency of occurrence (about 500,000 years once) and thus the hazards they cause should be studied with more care. The damage will be multifarious and far-reaching, ranging from the shock wave at the instant of impact to possible 'nuclear winter' and greenhouse effect years later. We hereby classify the effects of an impactor into three phases according to the different time they dominate:



● Direct Impact Phase (Short-term effect, last several hours to a few days)

In this phase, the impactor approaches the surface of the Earth like a fireball because vast amount of heat is produced by friction when it penetrates the atmosphere. Small asteroid is disrupted before it reaches the surface while larger one probably makes a crater. Shock wave and local earthquake follows and a great tsunami is inevitable if the target area is in the ocean. Nuclear Simulation Model can be used to estimate the size of crater according the different properties of the impactor. The damage is local in this phase.

● Nuclear Winter Phase(Mid-term effect, last several months to more than a year)

Large crater-forming impact ejects evenly-dispersed dust into the atmosphere and cuts off the sunlight. The dwindling solar energy, together with the subsequent drop of temperature, will threaten the food production. Other effects such as acid rain and snowstorm are also possible. We name it 'nuclear winter' after the similar effect of a global nuclear war. The damage is global in this phase and we develop Ecosystem Model to simulate the possible effect ton the ecosystem.

● Greenhouse Effect Phase (Long-term effect, last at least decades)

Months later, after the atmosphere is cleared, water vapor and carbon dioxide released to the stratosphere will produce an enhanced greenhouse effect. Thus, temperature will rise and the warming climate is likely to last for a long time. The global sea-level rise will ensue since the ice sheet in Antarctica melts. Sea Level Model is applied to estimate the change of temperature and the sea level.

Not all the NEOs will produce the damage enumerated above, impactors of small size and velocity makes only local damage. However, large asteroids of several kilometres in diameter are very likely to produce most of the damages. The problem of threshold condition between global catastrophe and local damage is discussed carefully in our paper.

Assumptions

● The shape of the asteroid is presumed to be a sphere and the density is the even in the interior.

● The landform of the South Pole is thought to be constituted of upper layer of ice sheet with 2000m in thickness and bottom layer of rock with density of $2.6g/cm^3$.

● The effect of atmospheric friction to the size and velocity of the asteroid is neglected. This can be justified by the scientific research outcome that all impactors except the large one will be decelerated in atmosphere to a nearly free-fall speed. The reduction of size from friction is typically tens of meters in diameter and is negligible compared with the diameter of our asteroid.

● The dust ejected from a crater is proportional to its volume. A parameter β is set as the ratio of the dust mass to the rock mass.

Models

For the reason of multifarious consequences and unprecedented severe damage, no single model can be used to deal with all the possible consequences of the impact. Because no disaster except nuclear war is comparable to the impact, our basic scheme is to use the simulation of nuclear war to assess the direct damage of the impact. As to long-term consequences, Ecosystem Model and Sea Level Model are developed to estimate the damage in their respective category.

Notation Table

d_0	initial diameter of the asteroid (km)
v_0	initial velocity of the asteroid (km/s)
m_0	initial mass of the asteroid (ton)
d_1	diameter of the asteroid after penetrating the ice sheet (km)
v_1	velocity of the asteroid after penetrating the ice sheet (km/s)
m_1	mass of the asteroid after penetrating the ice sheet (ton)
D	diameter of the crater (km)
H	depth of the crater (km)
ρ_{ast}	density of the asteroid (g/cm^3)
$\rho_i = 0.9$	density of the ice sheel(g/cm^3)
$\rho_r = 2.6$	density of the rock (g/cm^3)



Crater Formation on South Pole

Formation of the Crater

On impact, the bulk of the kinetic energy of the projectile is transferred to the land surface in the form of shock wave, which impacts the rocks, ice and projectile itself. The advancing wave front strongly compresses the materials around the target area. A varying fraction of them are vaporized, melted or fragmented by the high pressure and temperature. The initially intense shock diminishes with increasing radial distance from the line of penetration. Particularly, at the South Pole, upper layer of ice is likely to be vaporized and penetrated more rapidly because of its boiling point and low density.

Size of the Crater at the South Pole

Because the properties (diameter, velocity, density) of asteroids can vary continuously in a large scale, we convert the problem into a discrete one by choosing four asteroids A, B, C and D as our samples:

Tab.1 Four sample asteroids

Sample	d_0(km)	v_0(km/s)	ρ_{ast}(g/cm^3)	M(ton)	$E(mv^2/2)$(MT)
A	1	11.2	3	$1.57 * 10^9$	$2.3 * 10^4$
$B(B*)$	1	20	3	$1.57 * 10^9$	$7.5 * 10^4$
C	2	20	3	$1.257 * 10^{10}$	$6.0 * 10^5$
D	5	20	3	$1.96 * 10^{11}$	$9.4 * 10^6$

(1MT = 1megaton TNT = $4.168 * 10^{15}$J)

We choose 11.2 km/s as the velocity of A since it's the lower velocity limit of all the large asteroids if atmospheric friction is omitted. (see Appendix A). 3g/cm^3 is the mean density of the stony asteroid because most asteroids are made of stone.

Note here we have an additional sample B^*, which has the identical initial condition as B except that it hits on rocky surface instead of ice sheet. Through B^*, we can further discuss the different effect produced by Antarctic impact and impact elsewhere.

We get two formulas at hand to calculate the size of the crater:



①Using the relation between crater diameter and impact energy:

$D \approx 0.1\sqrt[3]{E}$ (Dence, Grieve and Robertson) (1)

D——the diameter of the crater(km);

E——kinetic energy of the asteroid (1000tTNT)

②The formula based on the studies of cratering in nuclear tests:

$D = S_1 * S_2 * C * K * W$ (Gene Shoemaker) (2)

$S_1 = (g_0/g)$——gravity correction factor for bodies other than Earth, where $g_0 = 9.8$ m/s∧2 and g is the surface gravity of the target body.

$S_2 = (\rho_0/\rho)$——correction factor for target density $\rho_0 = 1.8$ g/cm∧3, ρ is the density of the target material.

C: crater collapse factor, 1 for craters ≤ 3 km in diameter, 1.3 for larger craters.

K: 0.074 km/(kT TNT equivalent) empirically determined from nuclear test crater.

$W = Pi * d_0 * \rho_{ast} * v_0/(12 \times 4.185 \times 10)$ projectile kinetic energy of TNT equivalent especially for the impact on Earth, we have:

$$D = \frac{d_0 * \rho_{ast} * v_0 * C}{\rho} \times 8.334 \times 10^{-4}$$

$H = (1/5 \sim 1/6) * D$ (Elbert. A. King, [5])

Of the two models mentioned above, we choose the latter because it takes more factors into consideration, including the size, velocity and density of the asteroid as well as the properties of the target area.

Another problem is the two-layer structure of the landform in Antarctica. It is quite clear that the crater in ice sheet is of much larger size compared with the crater in rock due to their different density. What will happen if the asteroid penetrates the ice sheet first and then hits the rock? We prompt the following method to solve this problem:

Two further assumptions are made:

● The shape of the crater is roughly regarded as a cone with diameter/depth = 5.5/1;

● The decrease of both the mass and velocity of the asteroid is proportional to the depth it has penetrated.

Three steps are taken to calculate the size of the crater:

Step 1: Assume the asteroid hits on the pure ice surface and makes an ice crater. The crater depth and diameter can be calculated from (2)

Step 2: According to further assumption, the remains of the velocity and mass of the asteroid after it penetrates the 2km ice sheet are calculated as follows.

$$\frac{m_1}{m_0}=\frac{v_1}{v_0}=\left(\frac{d_1}{d_0}\right)^3=\frac{H-2(\mathrm{km})}{H} \tag{3}$$

Step 3: With the remained mass and velocity calculated in the last step, calculate the size of the rock crater below the ice sheet from (2)

The three steps can be summarized to the following equation set:

$$\begin{cases} H_{ice}=\dfrac{D_{ice}}{5.5}=\dfrac{d_0 * \rho_a * v_0 * C}{5.5\times\rho_{ice}}\times 8.334\times 10^{-4} \\ \dfrac{v_1}{v_0}=\left(\dfrac{d_1}{d_0}\right)^3=\dfrac{H_{ice}-2(\mathrm{km})}{H_{ice}} \\ H_{rock}=\dfrac{D_{rock}}{5.5}=\dfrac{d_1 * \rho_{ast} * v_1 * C}{5.5\times\rho_{rock}}\times 8.334\times 10^{-4} \end{cases} \tag{4}$$

Calculate the crater size of four sample asteroids using (4), we have:

Tab.2 Crater size comparison

Sample	D_{rock}(km) on Antarctica	D_{rock}(km) on other Continent	D_{ice}(km) of ice crater	Percent of rock destroyed
A	9.2	14.0	40.4	28.1%
B(B*)	20.1	25.0	72.2	51.7%
C	46.0	50.0	144.5	77.9%
D	120.0	125.0	361.1	88.5%

The last column 'percent of rock destroyed' shows the ratio of rock crater volume(under ice sheet) on Antarctica to the rock crater volume on other continent without ice sheet. The asteroid energy is diminished by the ice sheet and creates relatively small rock crater. Small impactor is influenced by ice sheet greatly (such as A, which only destroy 28.1% of the rock after penetrating the ice), while large

impactor (like D) nearly never 'see' the ice sheet and create the crater as big as usual. This phenomemon should be paid more attention for it displays the difference in consequences between Antarctic impact and impact elsewhere.

Nuclear Winter and Food Production

According to the analysis above, nuclear winter caused by dust will become the main killer of the ecosystem. So we particularly concentrate on the emission of dust and its influence on the solar energy and temperature. The dust ejected to the atmosphere can be estimated from the calculated volume of crater. But we haven't determine the emission ratio β.

In our calculation, we estimate the value of emissioon ratio β to 1.5% from the data of nuclear test (see Appendix B). Thus the dust ejected by the four sample asteroids can be computed:

$$M_{dust} = \beta * M_{rock} = 1.5\% * V_{crater} * \rho_{rock} \tag{5}$$

The results are shown in the following table:

Tab.3 Dust emission

Sample	$V_{crater}(m^3)$	M_{rock}(ton)	M_{dust}(ton)
A	$3.67 * 10^{10}$	$9.54 * 10^{10}$	$1.24 * 10^{9}$
B	$3.95 * 10^{11}$	$1.03 * 10^{12}$	$1.34 * 10^{10}$
B*	$7.44 * 10^{11}$	$1.93 * 10^{12}$	$2.51 * 10^{10}$
C	$4.63 * 10^{12}$	$1.21 * 10^{13}$	$1.57 * 10^{11}$
D	$8.23 * 10^{13}$	$2.14 * 10^{14}$	$2.78 * 10^{12}$

Two things should be emphasized from the table above:

First, the dust emissions of the four sample asteroids seem to differ with eath other in the order of the number. That is to say, the smaller one emits approximately 1/10 dust of the larger one although they are very close in diameter and velocity. This phenomemon can be explained like this: From (2) the following relation is drawn:

$$M_{dust} \propto V_{crater} \propto D^3 \propto v^3 d^3$$

Thus, an asteroid with two times the diameter or velocity of another one is



likely to create a crater $2^3 = 8$ times larger. Furthermore, it will eject 8 times of dust.

Second, sample B^* hits on land ejects two times dust of sample B hits on Antarctica (compare B^* and B). This dicrepency will show its significance in our later calculation.

Solar Energy

The dust will be ejected into atmosphere in a few weeks, after that it will slowly drop back to ground in a long period ranging from several months to years. During this period, solar energy will fall in a large scale (nearly to zero) rapidly at beginning and then slowly recover to its normal value. We simulate this solar energy change by two steps:

Step 1: Any single particle of dust is likely to fall at any time with the same possibility. That is to say the dust-deposit speed is proportional to the amount of dust in atmosphere:

$$\frac{\mathrm{d}M_{dust}}{\mathrm{d}t} = \lambda_1 M_{dust}$$

where, λ_1 is a constant.

thus, $M_{dust} = c_1 e^{-\lambda_1 t}$

Through boundary condition $M_{dust} = M_0$ at $t = 0$, we have $c_1 = M_0$ (M_0 is initial dust mass)

So final equation is $M_{dust} = M_0 e^{-\lambda_1 t}$ (6)

Step 2: A thin layer of dust of mass dM will absorb a certain ratio of solar energy reaching its upper surface (see Fig. 1). The following relation is drawn:

$\frac{\mathrm{d}E}{\mathrm{d}M_{dust}} = \lambda_2 E$, where λ_2 is another constant

Similarly we have

$$E = E_0 e^{-\lambda_2 M_{dust}} \quad (7)$$

if E_0 is the normal solar energy.

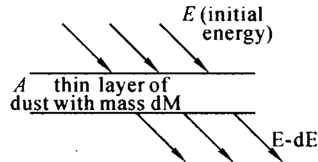


Fig. 1 Solar energy cut-off

Connect (6) and (7) we have: $E = E_0 e^{-\lambda_2 M_0 e^{\lambda_1 t}}$ (8)

Although the constants λ_1 and λ_2 are hard to determine, the tendency of the solar energy variance can be ploted according to Equ. 3 using Mathematica:

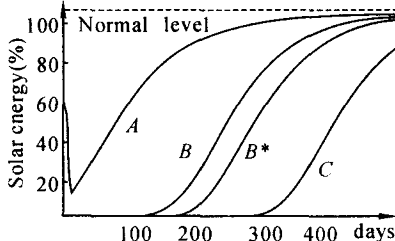


Fig. 2 The solar energy variance vs time

From the figure presented above, the dust accumulates quickly in the atmosphere which is represented by the 'quick fall' in the beginning of the curve, after that it dissipates in a long period ranging from months to years. Larger asteroid's dust cuts off more sunlight and the process of dissipation is longer. of the four samples, only A doesn't cut off all the sunlight while the other three seems to keep the whole world in darkness for hundreds of days. There's no doubt that B* causes more severe effect to solar energy than B does.

Food Production in Southern Hemisphere

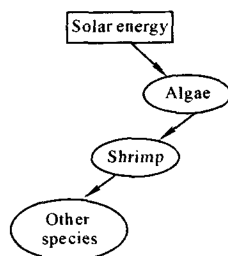
As to plants, food production is decided by the difference between the function of photosynthesis and prespiration. Insufficient solar energy will decrease or stop photosynthesis. Photosynthesis will no longer dominate in plant, resulting in decreased food production. Then, in turn, food production of animals reduces since they take plants as their food. This process can be simulate by Volterra's predator-prey model.

Here we choose two typical species in southern hemisphere from the bottom of the food chain as the objects of our analysis: algae and shrimp which both have crucial effect to food production. Algae lives on solar energy while shrimp lives on algae. Other marine species will also affected by them since they are in the same food chain.

Predator-prey model:

$$\frac{dx}{dt} = x\left(r_1 - \frac{x}{N_1} - \sigma_1 \frac{y}{N_2}\right)$$

$$\frac{dy}{dt} = y\left(-r_2 + \sigma_2 \frac{x}{N_1} - \frac{y}{N_2}\right) \quad (9)$$



x: algae production

y: shrimp production

σ_1: capture ability parameter of shrimp

N_1: resource limition of algae

N_2: resource limition of shrimp

r_1: reproduction rate of algae

r_2: death rate of shrimp

σ_2: food provision parameter by algae

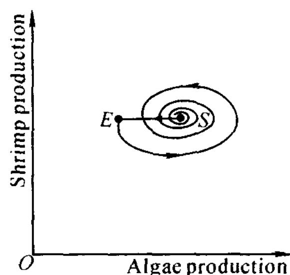


Fig.3 Sample A: small fluctuation

Note that here we regard x as the production of algae instead of its number, so it's reasonable to suppose production rate r_1 proportional to the solar energy it absorbs. That's because the function of photosynthesis is proportional to solar energy with low light intensity and the food production is proportional to intensity of photosynthesis. Thus we substitute r_1 by λE. In fact, the accurate value of other parameters in Volterra's model should be determined by the knowledge of biology which is beyond our scope. However, the tendency of food production change can still be learned. Because Volterra's model has no analytic solution, we work it our numerically by programming and plot the curve of both algae and shrimp productions as follows:

Analysis:

- When the solar energy dwindles quickly at beginning, the algae production also decreases rapidly from stable point S to E which causes a decline of shrimp production. While the solar energy slowly returns back to its normal level, the production of either algae or shrimp gradually returns to S through a set of spiral route instead of direct return.
- Pay attention to the figure at the right column (Fig.4 and Fig.6): the production

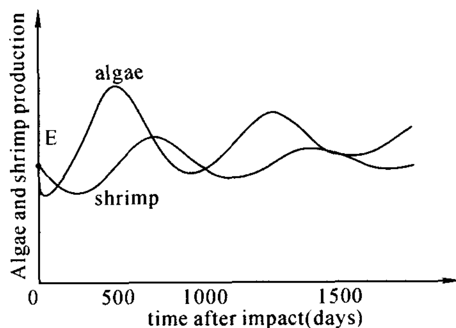


Fig.4 Sample A:mid-term effect

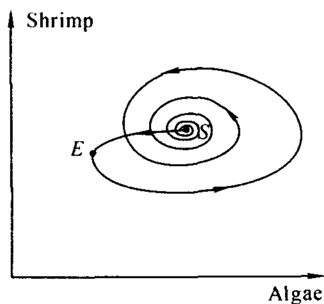


Fig.5 Sample B:large fluctuation

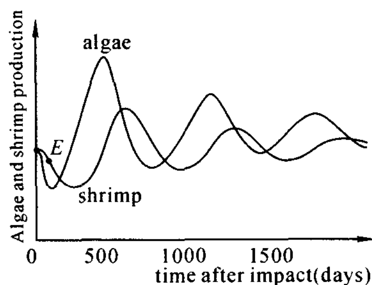


Fig.6 Sample B:long-term effect

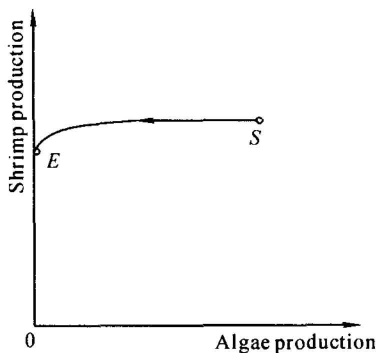


Fig.7 Sample B^* :destruction of ecosystem

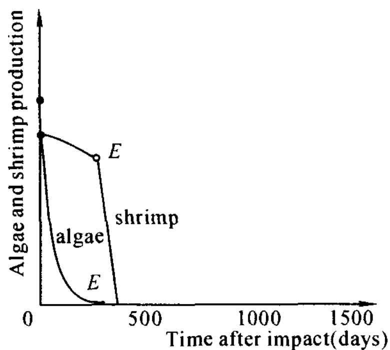


Fig.8 Sample B^* :quick extinction



of shrimp doesn't react immediately to the change of the algae production. Instead, there is a 'phase laggard' of 1/4T between the change of algae and shrimp.

- Compare the resultant curve of A and B with that of B*, we find a very obvious threshold between total annihilation of the ecosystem and fluctuation of food production. If the insufficiency of solar energy is too serious, the algae production can't 'rebound' back to its normal level. Such as sample B*, algae production reduced to zero because of long-term absence of sunlight, which in turn, kills all the shrimps and inevitably leads to extinction of other species. The impact greater than B^* (say C or D) will have the same result. In contrast, if only the algae can survive this ordeal at beginning (like A), the ecosystem will come back to norm.
- Even if the algae and shrimp survive the initial period of time, their production will 'oscillate' violently for a long period. The greater the impact is, the larger the amplitude of fluctuation is and the longer it will last. (see following table)

Tab.4 Recovery Speed

amplitude as the percent of norm(algae)	30%	20%
Recovery time (days) by A	600	1300
Recovery time (days) by B	1100	1450

Sea Level Change

The estimate of sea level is quite out of our expectation!

The following factors are supposed to have their respective effects to sea level:

- Ice melting by the direct impact of asteroid.
- Vaporization of the sea water.
- Ice melting by the long-term change of the seawater temperature.

At first, the upper limit of the amount of melting ice by direct impact is calculated. The energy used to melt ice is estimated by the energy of breaking the ice sheet. Volume of water (0℃)melted from ice sheet with mean temperature −40℃ are calculated:

Tab.5 Sea level rise by direct impact

sample	A	B	C	D
Volume of water melted from ice(km^3)	140.79	292.4	1260.0	8208.9
sea level rise (mm)	0.39	0.81	3.49	22.7

Other two factors are considered as follows:

The atmospheric temperature is subject to the CO_2 density in atmosphere as well as the dust ejected in atmosphere. The following regression equation (r. Bryson, 1970) is used to describe this relation:

$$\Delta T = -3.546 + 0.012A - 0.002B - 0.006C$$

where A is CO_2 density (ppm), B is dust density (mg/1) and C is sunspot constant. The temperature before impact nearly remains stable, thus

$$\Delta T_0 = -3.546 + 0.012A_0 - 0.002B_0 + 0.006C_0 = 0$$

If ΔA and ΔB are the change of CO_2 and dust density due to impact, we get:

$$\Delta T = -3.546 + 0.012(A_0 + \Delta A) - 0.002(B_0 + \Delta B) + 0.006C_0$$
$$= 0.012\Delta A - 0.002\Delta B \quad (10)$$

Till now we have three relations uncertain:

- Relation between the atmospheric temperature and amount of vaporization.
- Relation between variance ratio of ocean temperature and difference between ocean and atmosphere temperature.
- Relation between ice melting speed and ocean temperature.

Because we don't know the exact relations mentioned above, here we follow the engineers' advice: Using linear relation where you don't know the exact relation. According this assumption, the following complicated equation set is drawn:

$$
\begin{cases}
T_{air} = T_0 + 0.012\int_{t_0}^{t}\Delta A \mathrm{d}t - 0.002\int_{t_0}^{t}\Delta B \mathrm{d}t \\
\dfrac{\mathrm{d}T_{ocean}}{\mathrm{d}t} = \lambda_1(T_{air} - T_{ocean}) \\
V_{vapor} = \lambda_2 T_{air} \\
V_{melt} = \lambda_3 T_{ocean} \\
\dfrac{V_{melt} - V_{vapor}}{S_{ocean}} = h_{rise}
\end{cases}
\tag{11}
$$

where the amount of ΔA, ΔB can be calculated from the crater size; S_{ocean} is the area of the ocean and h_{rise} is the sea level rise. But we still have problem of determining the value of λ_1 λ_2 λ_3, which, as mentioned before, can't be determined reasonably. So we just simulate the tendency of sea level instead of precise value using the (11).

(The simulation program is designed in C and runs in MS－Windows)

From Fig. 9～Fig. 11, three phases are very obvious in both temperature and sea level change. Surprisingly, the sea level doesn't rise continuously because of melting ice as we expected before. Instead, the sea level undergoes the first rise, fall and then rise again (totally three phases). It's unbelievable at first glance, but our analysis comes later explains it well. Please note that the sea level change has been exaggerated in Fig. 11.

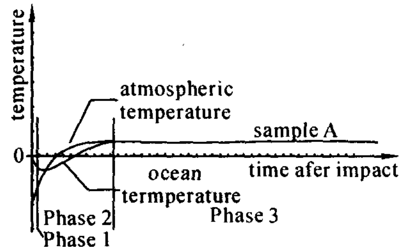


Fig. 9 Tempertuare variance vs time simulation Sample A

Analysis:

Before analysis, one thing should be pointed out: The stationary state of sea level is a dynamic balance between the function of ice melting which raises the sea level and vaporization which lowers the sea level. Ice melting is mostly affected by ocean temperature while vaporization is affected by atmospheric temperature.

Phase 1: The atmospheric temperature drops rapidly for absence of sunlight, but ocean temperature 'reacts' slowly and almost remains stable. So the vaporization function reduces while ice melting continues, results in sea level rise. The starting points S_A, S_B are higher than normal value because of the melting ice by direct impact.

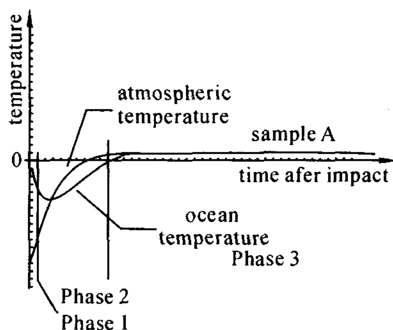


Fig. 10 Tempertuare variance vs time simulation Sample B

Phase 2: In this phase, ocean temperature 'catches up with' the step of atmospheric temperature and fall in a large scale. Meanwhile, atmospheric temperature comes back towards norm because dust seems to dissipate. This time, vaporization dominates, lowering the sea level.

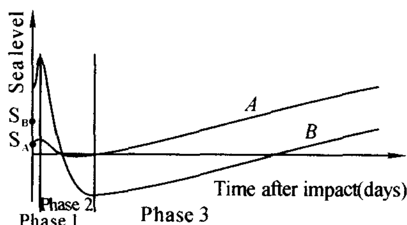


Fig. 11 Sea level change vs time simulation Sample A and B

Phase 3: The dust has been completely dissipated, both ocean and atmospheric temperature recover to normal values. The greenhouse gases (say CO_2) emitted during impact begins to dominates. It causes the temperature a little higher than norm, which leads to continuous melting of ice sheet and the long-term sea level rise.

We conclude as follows: After the impact of asteroid, the world temperature of both atmosphere and ocean will be higher than before, and sea level seems to rise slowly if no other factors stop it.

Amount and location of human casualties

Since Antarctica is nearly an unpopulated area except some scientific stations, direct impact of the asteroid(crater and shock wave) is unlikely to cause any consid-

erable human casualties. So we turn to other continents: South America, Africa, Australia and New Zealand because they're the closest lands from Antarctica. Two possible types of disaster are probed: earthquake and tsunami. The earthquake in Antarctica has little chance to pass to other continents for the South Pole is right at the middle of Antarctic plate. As to tsunami, it is more likely to be triggered by the earthquake at the bottom of the sea, so the incidence is also quite low. Even if we know tsunami or earthquake will happen, its resultant damage is impossible to be estimated.

However, potential danger brought about by the change of sea-level and food production can't be neglected. Through our computer simulation, any asteroid larger than 1 km in diameter will eject enough dust to cut off sunlight for several months. The food production will decrease violently which causes severe shortage of food supply. The following table suggests us the possible location of the famine:

Tab.6 The world's food supply

Location	Shortage		Severe shortage	
	Percent of Population	Population (million)	Percent of Population	Population (million)
South Africa	44	150	25	90
East Asia & Oceania	14	40	7	20
South Asia	50	470	21	200
Middle East & North Africa	10	20	4	10
South America	13	50	8	20

—World Bank, Poverty and Hunger, 1986

This global scale famine will cause starvation of at least 730 million, 340 million among which are almost destined to death in this disaster. The famine is likely to happen in the developing countries and poor area according to their present food supply situation, such as South Africa and South Asia. However, if the disaster lasts longer, the famine will endanger even larger population, including Europe and North America.

What's the difference to choose Antarctica as a victim

This problem is of essential importance in our analysis. According to all the calculations, we learn that it will be a great difference between Antarctica impact and impact elsewhere. Three aspects are discussed:

Direct impact

Direct impact by asteroid seems to create less damage in Antarctica than other places in the world since it's unpopulated and has no cities. Otherwise, it will produce great tsunami if hits the ocean or terrible earthquake if hits the land, human death of large scale seems to be inevitable in both situations.

Long-term effect

Different craters created by asteroids in Antarctica and other lands will result in different long-term consequences due to the amount of dust they eject. Our contrast experiment between B and B^* gives us an estimated magnitude of this difference. B^* creates a crater 25% bigger than B in diameter (see Tab.2), ejects nearly two times the dust of B(see Tab.3) and affects the solar energy in different degree(see Fig.2). If the threshold of ecosystem survival is right between B and B^* (It's quite possible due to Fig. 3 ~ Fig. 8, although these figures just show a tendency), Antarctica will save us from extinction while other continents will put us at risk. Anyway, disaster caused by Antarctica impact is significantly smaller.

As to sea level rise, from Fig.11, we know it is mainly caused by the greenhouse effect in long-term instead of melting ice by the direct impact. Greenhouse effect is the common characteristic of all the large scaled impact in spite of the target area location, so we can't say Antarctica impact raises sea level more as we expected before.

Possible inverse of geomagnetism

We make this supposition from our everyday experience: If a small magnet is hit violently enough by something else, it loses its magnetism. Asteroid impact



probably has the similar effect to geomagnetism because the South Pole is very close to magnetic south pole of the Earth. Another factor supports our supposition is the scientists' discovery: geomagnetism has inverted several times in history.

If our supposition is true, severe consequence ensues: Since geomagnetism shields us from the cosmic rays, loss the magnetic mask during the inverse of geomagnetism will expose the whole ecosystem to the harmful cosmic rays. The extent of damage is unknown to us.

Threshold and parameter sensitivity

Since it's hard to know the accurate damage, we adopt a widely accepted criterion to estimate the threshold condition of global catastrophe. The dust needed to cause a severe nuclear winter is $10000\text{Tg} = 1 * 10^{10}$ ton. Of our five samples, only A is below the threshold while other four exceeds. So all the samples except A are likely to make global catastrophe.

Three 'threshold asteroids' are given as following:

If $v = 11.2\text{km/s}$, $d = 2.334\text{km}$

If $v = 20\text{km/s}$, $d = 0.933\text{km}$

If $v = 30\text{km/s}$, $d = 0.871\text{km}$

Because they approach the Earth at average 20km/s, the asteroids having diameter larger than 1km is likely to lead to global catastrophe.

The three-dimensional graph of dust emission versus velocity and diameter is shown:

For dust amount is used as the criterion of disaster, we calculate the variance ratio of M_{dust} versus the change of d and v. The analytic solution can be drawn from (4) and (5):

$$\frac{dM_{dust}}{dv} = \frac{1.17 \times 10^{-5}\pi(-9.9 + 3.25 \times 10^{-3}dv)^3}{v} - \frac{8.97 \times 10^{-4}\pi(-9.9 + 3.25 \times 10^{-3}dv)^4}{dv^2} \tag{12}$$

$$\frac{dM_{dust}}{dd} = \frac{1.17 \times 10^{-5}\pi(-9.9 + 3.25 \times 10^{-3}dv)^3}{d}$$

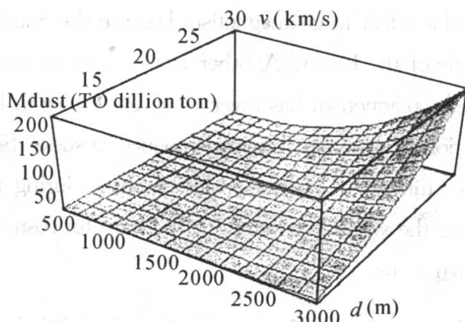


Fig.12 Amount of dust ejected versus velocity and diameter of the asteroid

$$-\frac{8.97\times10^{-4}\pi(-9.9+3.25\times10^{-3}dv)^4}{d^2v} \tag{13}$$

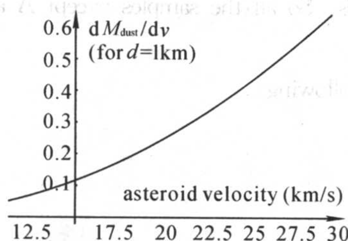


Fig. 13 Dust amount variance rate versus velocity

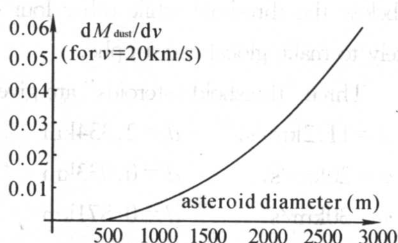


Fig.14 Dist amount variance rate versus diameter

Using the equation presented above, we plot the following graph to learn the amount of dust variance rate versus the velocity or diameter of asteroid:

From the two figures above, we find the dust variance rate becomes larger when the diameter or the velocity of the asteroid is larger. That is to say, if the asteroid is very large or rapid, a small change add t_0 its velocity or diameter will contribute a lot to the dust it ejects. Consequently, the damage will be much more serious. But when the asteroid is small and slow (say, 11.2km/s with 0.5km diameter), the velocity or diameter change of the same amount has little contribution to the dust amount.



Strength and Weakness

Strength

- Our computer model of simulation gives the tendency of the following changes:

—Solar energy change with time

—Food production change with time

—Temperature variance of both atmosphere and ocean to ward time

—Sea level variance with time

These tendencies are all shown in form of graph, which is a very intuitive way to organize data, Moreover, we plot the curve of asteroids of different properties in our graphs which enable us to find out the relation between their size and their consequences.

- We analyze the difference between Antarctica impact and impact on other continents through our contrast experiment of B and B^* .
- We pay enough attention to the special landform of Antarctica, two-layer structure of landform is assumed in calculating the crater size.
- Our computer program of food production contains many biological constants (such as death rate, reproduction rate) that are easy to adjust. So the model can be used in different surroundings if these constants are known.
- We include many factors in our model of sea level variance, including vaporization speed, ocean temperature, CO_2 and dust density, etc. If their weight in our model is set properly, our estimate will be more reliable and comprehensive.
- Threshold condition between local damage and global disaster is estimated and parameter sensitivity is tested.

Weakness

- In our computer simulation, the exact result can't be given because many parameters in our model remain uncertain. So accurate estimate can't be made, only tendency and basic rules are presented.
- No real data available to test our model.

● The three linear relations assumed in our model of sea level may be inaccurate. However, we believe linear relation holds approximately in a small range.

● Our assumption that the crater is nearly a cone is too simple.

● The assumption that the emission of CO_2 and dust are proportional to crater size may be inaccurate.

● The assumption of the relation between the solar energy and dust density as well as the dust dissipation process can be questioned.

● Using the empirical formula of nuclear explosion and regression equation of temperature variance makes our model vulnerable to error and prejudice of others.

Conclusion

Our model allows us to estimate the size of the crater, given the diameter, velocity and density of the asteroid. The threshold asteroid diameter between local damage and global catastrophe is about 1km if mean velocity is 20km/s.

The impact has far-reaching effect on ecosystem, the world's temperature, sea level, and consequent threat to human being. The stronger the impact is, the deeper the effect will be.

The impact on Antarctica causes less damage than impact on other continents and produces trivial direct human casualties.

Appendix A: Computation of the lower limit of asteroid velocity

The lower limit of velocity occurs when the asteroid with initial velocity zero is captured by the Earth's gravity filed. Let E_k represent the kinetic energy of the asteroid before it hit the surface, E_h represent the potential energy. We have:

$$E_k = \frac{mv_{\max}^2}{2} = -E_h = -\int_{R_0}^{\infty} \frac{GM_e m}{R^2} dR = \frac{GM_e m}{R_0}$$

where M_e is the mass of the Earth, m is the mass of the asteroid and R_0 is the radius of the Earth.

Thus, $$v_{\min} = \sqrt{\frac{2GM_e}{R_0}} = 11.2(\text{km/s})$$

Appendix B: Estimate of the dust emission ratio from nuclear test



In nuclear tests, the following empirical relation can be drawn[1]:

Type of burst	Dust mass(ton/MT)
L'and surface	$3.3 * 10^5$
Land near-surface	$1 * 10^5$

For asteroid will impact on the surface of the Earth, it is reasonable to choose 'land surface t', and the TNT equivalent energy can be calculated from $E = 1/2 * mv^2$:

Dust amount $= 3.3 * 10^5 * mv^2/2$

For sample B, its kinetic energy is $7.5 * 10^4$ MT, so the dust it ejects is $2.475 * 10^{10}$(ton)from the equation above. If B hits the land, it makes crater of diameter 25km and height 4.5km according to the table. The total mass of the rock in crater is $1.934 * 10^{12}$(ton). So $2.475 * 10^{10}/1.934 * 10^{12} = 1.3\%$ of the rock is ejected to atmosphere in the form of dust. We also test sample A,C,D using nuclear explosion data, similar ratio can be drawn.

References

[1] R.P.Turco, O.B.Toon, T.P.Ackerman, J.B.Pollack, Carl Sagan: Nuclear Winter: Global Consequences of Multiple Nuclear Explosions. Science V 222: 1283 - 1292(1983)

[2] Clark R.Chapman &David Morrison: Impacts on the Earth by asteroids and comets: assessing the hazard. Nature.V 367:33 - 39(1994)

[3] Billy P. Class: Introduction to Planetary Geology. Cambridge University Press 1982

[4] Halliday, Resnick, Krane:Physics

[5] Elbert A.King:Space Geology an Introduction 1976

[6] Norman H.Sleep, Kevin J.Zahnle, James F.Kasting & Harold J.Morowitz: Annihilation of ecosystems by large asteroid impacts on the early Earth. Nature V 342:139 - 142(1989)

[7] Christopher F.Chyba, Paul J.Thomas & Kevin J.Zahnle: The 1908 Tunguska explosion:atmospheric disruption of a stony asteroid. Nature V 361:40 - 44

(1993)
[8] Encyclopedia Britannica 1998
[9] Robert E. Norris World Regional Geography
[10] Rabinowitz, D. L. et al: Evidence for a near-Earth asteroid belt. Nature 363: 704 - 706(1993)
[11] Chyba, C. F.: Explosions of small Spacewatch objects in the Earth's atmosphere. Nature 363: 701 - 702(1993)

点 评

在竞赛期间,本文作者仅用了不到一天的时间,从网上查阅了有关陨石坑、地震、海啸、南极地区地貌气候及生态情况等方面的大批资料,阅读了发表在 Nature 等杂志上的 200 多页的相关文献。在此基础上研究了小行星撞击冰层、岩石后可能引起的灾难情况,包括冰层融化造成的海平面上升、飞扬的尘土造成的大气温室效应、南极地区地理气候及生态环境的变化情况、对南极海域海藻磷虾及其他生物成长的影响,等等问题。作者考虑问题较为细致,广泛地应用了已有的研究结果,如有关陨石坑的资料、相关的地理地质生态资料、食饵 - 捕食者系统的数学模型(即 $P-P$ 模型)等,使文章显得较为殷实生动,具有较强的说服力。由于小行星撞击南极事实上并未真正发生过,而且不同的小行星具有不同的质量、速度和密度,撞击南极会造成不同程度的灾难,作者们采用了引入参数的方法来研究不同情况下可能产生的后果,既巧妙地克服了缺乏真实数据的困难,又在一定程度上完成了灵敏度分析。本文组织得较好,条理也十分清楚,在每一问题的研究之后,作者们还充分利用了图形、表格来加以总结,使研究结果显得更为明了清晰。

本文获得了 1999 年美国大学生数学建模竞赛一等奖。

Locate the pollution source

沈 权 杨振羽 何晓飞

Summary

We develop a model for the detecting strategy of new pollution. Among many factors affecting the movements of the pollutants, three processes govern the movements of pollutants in groundwater: advection, dispersion and retardation. Information from the wells is used to:

(1) Determine the rate and direction of groundwater movement.

(2) Determine the horizontal and vertical extent of the pollutants.

(3) Analyze the underground structure and characteristics.

Regarding the diversity and complexity of the given data, we employ a two-step data selection to determine the pollutants with the most possibility to cause new pollution during this period of time.

We refine the data to choose those representing the variation best during this period of time.

Then, by using the grid-search algorithm, we write a computer program to simulate the movement process and identify the location and start time of the pollution source. The program is written in C Language and runs on a Pentium 166.

Four kinds of new pollution source are located . The curve result from our model is in a good agreement with that of the given data.

Finally, parameter sensitivity is tested.

Assumptions

The development of the mathematical model requires several assumptions regarding the geological characteristics of the ground. The following are our assumptions:

(1) All soil and aquifer properties are homogeneous and isotropic throughout both saturated zone and unsaturated zone.

(2) The aquifer consists of sand and gravel.

(3) Steady, uniform water flow occurs only in vertical direction throughout the unsaturated zone, and only in the horizontal (longitudinal) plane in the saturated zone in the direction of groundwater velocity.

(4) The physical process plays greatest role while the chemical process is negligible.

(5) All the parameters describing the characteristic of both zones are constant throughout the monitoring period.

(6) All the sources of the pollutants are point sources.

Other assumptions are made during the course of developing the model.

Problem One

This problem is dedicated to estimate the location and start time of the source, therefore we have to consider the movement process of the pollution and the structure of the underground.

Data Analysis and Processing

Whereas the diversity and complexity, we process the data as follows. We assume that there is no interact between each two pollu-

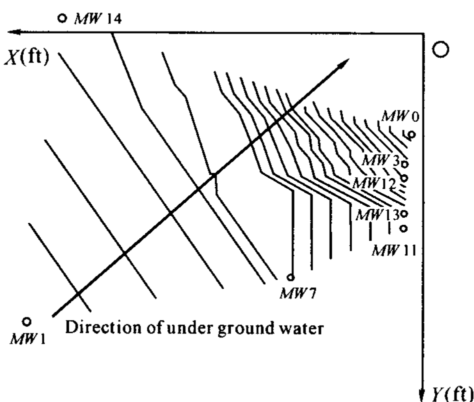


Fig. 1



tants. Thus, we can process each pollutant separately.

With the given data of the coordinate and water-level of each well, we plot the water-level map by using linear interpolation on the elevations of the monitoring wells, as in Fig1. According to the figure given above, for simplicity of computation, we assume that all the underground water flows in the same direction.

Data selection

Because we have thousands of data about concentration of varies of pollutants. It's necessary for us to select data correctly, otherwise they will confuse us and cause trouble in solving this problem. We do this work in the following steps:

● Because pollutants are strongly influenced by layers of different permeability, measurements of critical parameters and pollutant concentrations need to be done at intervals over the depth of the aquifer. The method for sampling at different depth in an aquifer is needed. By analyzing the data set, we find that almost each pollutant only affects one part of a certain well (top, middle, or bottom). For illustration, the changes of Calcium concentration is listed in Tab.1.

Tab.1

Monitor well	Concentration Change	Monitor well	Concentration Change
M1B	Always 26	M12B	Always 20
M3B	22→32	M12M	42→88
M3M	41→61	M13B	Always 23
M7B	Always 27	M14B	Always 33
M7M	50→115	M27B	Always 19
M9M	44→77	M33B	Always 20
M11B	Always 23	M33M	21→56
M11T	39→63		

Thus, for each pollutant we only need to consider the effect on one layer of the well. Furthermore, the data from the bottom of each well(if any) almost remain the same. Thus, we neglect the data from the bottom of the wells.

● We delete the data of some pollutant, such as Tetrachloroethane, Acrolein, Ben-

zene, Bromomethane, hlorobenzene, Cobalt, etc, whereas there are hardly any changes about these pollutants in each well.

● We think that the pulse fluctuation about the pollutant concentration during a relatively stable period, such as Manganese, is caused by random factors. Thus, we eliminate these pollutants from the data set.

● There is a particular constituent, the CarbonTotalOrganic, whose concentration value decreases significantly, from more than 1000 to less than 1.5. Thus, we eliminate it.

● Now, there are only four pollutants left. These are Calcium, Chloride, Magnesium and TDS.

Reselection

For a certain pollutant, in order to accurately reflect the tendency of the concentration change, we reselect its data as follows:

● Regarding the repetition of the data, for each well we choose two concentration value every year, one from the first half year and another from the later half year.

● Because we do not know the locations of MW − 27 and MW − 33, meanwhile, the concentration changes in the two wells are small, we do not consider these data.

● According to the groundwater flow direction, the average concentration value of MW − 9 should not be higher than that of MW − 3 and MW − 12 which contradicts the given data such as for Calcium, Chloride, etc. This is also true for Barium(In the year of 1997, concentrations in MW3M and MW12M vary from 50 to 85 while from 80 to 95 in MW9M). Therefore, we think that MW − 9 is a pumping well (Fig2). Thus, we do not use the data from MW − 9 in the following model analysis.

Finally, we list the data for Calcium that we will use to calculate the source location.

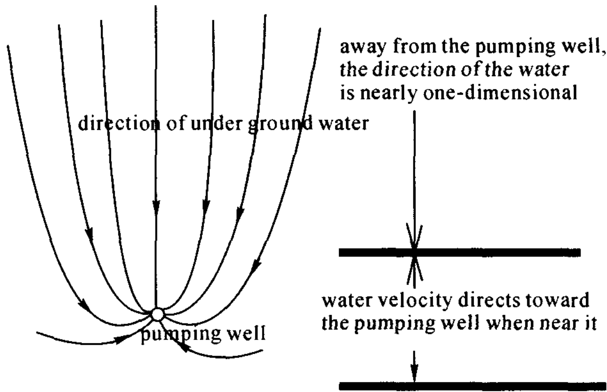


Fig. 2

Tab. 2

Date	MW - 3M	MW - 7M	MW - 11T	MW - 12M
12/7/93	41	50	39	42
3/7/94	42	50	43	47
9/19/94	42	45	41	41
7/10/95	36.5	54.3	44.7	59.5
10/10/95	19.2	53	43.2	54.7
3/6/96	62.4	65.1	50.7	82.4
10/9/96	60.2	61.9	53.3	87.6
3/18/97	63.8	125	53.2	87.6
12/15/97	61.4	115	63.8	88.4

●According to the given data, there is some pollutant detected in the early year such as 1990, we name it as the background concentration (C_b).

We think that the later pollutants concentration measured (as given data) consists of background concentration and new injected concentration. According to Fig. 1, MW - 1 must be at the headwater level. Moreover, the data from its bottom hardly changes during this period according to the data set. Thus, we estimate the value C_b by the data from MW - 1B as follows:



C_b = arithmetic mean of the concentration value from MW - 1B during this period for a certain pollutant

Symbols Used

C: Pollutant concentration (mg/liter)

C_b: Background concentration (described above)

R_d: Retardation factor

α_L: Horizontal dispersion coefficient

α_T: Vertical dispersion coefficient

V_d: Groundwater velocity (ft/day)

n: Effective porosity

t_0: The start time of the pollution (year)

q: The discharge rate of the pollutant (liter/day)

m: The discharge rate of the pollutant (mg/day)

C_0: The concentration in the pollutant source (mg/liter)

(x_0, y_0): The pollution source coordinate

θ: The angle between the direction of underground water and the x-axis.

Model Design

Model Formulation

Generally, the movement of pollutant consists of advection, dispersion and retardation. Furthermore, regarding the large scale of the area, the vertical movement is negligible. Thus, movement of pollutant in the soil (saturated and unsaturated) can be described by the following two-dimensional equation :

$$R_d \frac{\partial C}{\partial t} = V_d \alpha_L \frac{\partial^2 C}{\partial^2 x} + V_d \alpha_T \frac{\partial^2 C}{\partial^2 y} - V_d \frac{\partial C}{\partial x} \tag{1}$$

Model Explanation

Equation 1 is applicable under steady, uniform flow condition. Analyt-

ical solution to the movement equation can be developed for both continuous (step function) and pulsed inputs of pollutants as boundary condition. A step function implies the input of a constant concentration pollutant for an infinite amount of time, while a pulse load is a constant concentration input for a finite amount of time. The terms infinite and finite are relative to the time frame of the analysis.

According to the data set, we assume that the pollution source is applied as a step function (continuously) with the following boundary condition:

$$\left.\begin{aligned} &C(x,y,0)=0,\ (x,y)\neq(0,0) \\ &C(0,0,t)=C_0 \\ &C(\mp\infty,y,t)=0 \quad t\geqslant 0 \\ &C(x,\mp\infty,t)=0 \quad t\geqslant 0 \end{aligned}\right\}$$

Model Solution

The function is a second-order partial differential equation. Equations of this form apply to a wide variety of problems, including mass transport, fluid dynamics, and heat transfer.

For the case of an instantaneous point source at time t = 0, Eq1 possesses an analytical solution of the form

$$C(x,y,t)=S\cdot\exp\left(\frac{x}{2\alpha_L}\right)[W(0,b)-W(t,b)] \tag{2}$$

Where $m = C_0 q$

$$S=\frac{m}{4\pi V_d(\alpha_L\alpha_T)^{1/2}},$$

$W(u,b)$ is *hantush* − *function* described as follows:

$$W(u,b)=\int_u^\infty \frac{\exp[-y-b^2/(2y)]}{y}\mathrm{d}y$$

$$b=[x^2/(4\alpha_L^2)+y^2/(4\alpha_L\alpha_T)]^{1/2}$$

Before computing, we classified the parameter used according to our

assumptions above.

① During the data processing, the coordination and time of the pollution source are unknown, neither is the value m. Thus, x_0, y_0, t_0 and S are variable.

② The parameter α_L, α_T, θ, V_d are constants.

In this problem, the main task is to find the location and the start time of the pollutants. Here, we develop a grid-search optimization routine to get a optimized solution.

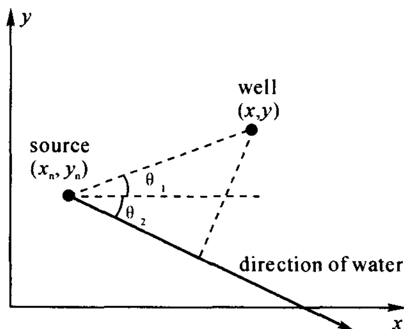


Fig.3

First, we approximately estimate the location of the pollutants. With the location, we transform the Cartesian coordinate as follows:

① Set the point of the pollutant source to be the origin point.

② Set the x-axis parallel to the direction of the underground water flow.

③ Set the y-axis vertical to the x-axis.

④ Coordinate-transformation:

$$R = [(x - x_0)^2 + (y - y_0)^2]^{1/2}$$

$$\theta_1 = \begin{cases} \sin^{-1}[(y - y_0)/R] & x \geqslant x_0 \\ \pi - \sin^{-1}[(y - y_0)/R] & x < x_0 \end{cases}$$

$$x' = R \cdot \cos(\theta + \theta_1)$$

$$y' = R \cdot \sin(\theta + \theta_1)$$

With the equation, the coordinate (x_i, y_i) of the ith well is transformed into (x_i', y_i').

Second, we construct an equation to calculate the movement of the pollutant under the ground. By substituting the well coordinate (x_i', y_i') into this equation, we calculate the concentration changes in each well and com-



pare with the changes according to the data set. We repeatedly adjust the location of the pollution source, the value S and the value t_0(detailed in the following) in grid, meanwhile, we reconstruct Cartesian coordinate until there is a satisfactory agreement.

The criterion for convergence is a simple sum of the square residuals between the data and the model predictions. The objective function to be minimized is as follows:

$$Y = \sum_i ((C_i - C_b) - C_i')^2 \qquad (i = \text{well number})$$

Where, C_i is the pollutant concentration data value for well i, C_i' is the value predicted by the model,

Where, C_i is the pollutant concentration value from the given data set, C_i' is the value predicted by the model, C_b is the background pollution level. and $C_i - C_b$ is the value for a new pollutant.

Parameter Estimation

The most important part of the modeling is estimation of reasonable and valid parameter values for a specific environment. We estimate the parameter of saturated zone as following,

(1) Hydraulic Conductivity K

Hydraulic Conductivity is considered only in the horizontal direction. According to the literature, we estimate $K = 265$ gpd /sq. ft.

(1 gpd/sq. ft. = $4.72 * 10^{-5}$ cm/sec)

(2) Hydraulic Gradient

According to Fig. 1 made by interpolation method, we assume that the direction of the underground water is one-dimension. With the equation of $I = \Delta H / \Delta L$, we can compute the approximate value $I = 0.01$ (refer to Fig. 1)

(3) Ground-water (Interstitial pore water) velocity V_d

According to Darcy's Law, V_d is defined as

$$V_d = -KI/n$$

Where I is the hydraulic gradient, K is the hydraulic conductivity, and n is effective porosity. We assume that the soil type of the saturated zone is sand with the porosity 20%, thus the estimate $V_d = 1.5$ft/day.

(4) Dispersion Coefficient α

This coefficient incorporate two forms of dispersive process, dynamic dispersion and molecular diffusion. According to the literature, the horizontal dispersion coefficient and the vertical dispersion are approximately equal. Both of them have the estimate value 25ft.

(5) Retardation Factor

Retardation is based on pollutant characteristics and aquifer composition. Since its affection is not very significant, we estimate a value of $R_d = 1$.

(6) Concentration in pollution source

According to the literature, the water table is usually sufficiently high so that the pollutant directly enters ground water, the C_0 value is the estimate of the source concentration.

Result

After processing the data by the computer program, we found that there are four new pollutants. These are Calcium, Chloride, Magnesium and TDS. Here, we present the location and the start time of the new pollution source predicted by our model as follows:

Tab.3

Pollutant	X-coordinate (feet)	Y-coordinate (feet)	Start time (m/d/y)
TDS	7077	6538	8/12/91
Magnesium	6423	7461	1/1/94
Chloride	6931	5823	5/18/91
Calcium	7750	6040	9/1/93

Finally we mimic the movement process of the pollutant reversely and compare with the given data set as following,

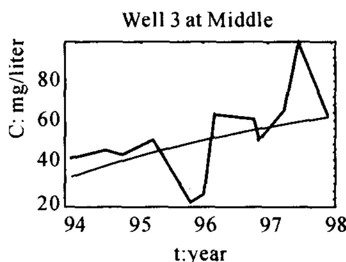


Fig.4

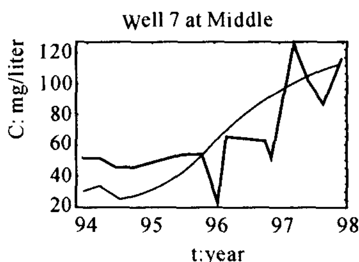


Fig.5

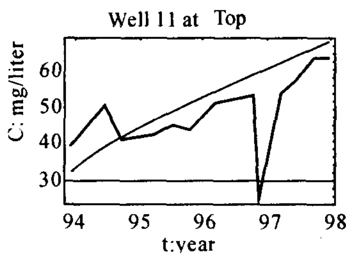


Fig.6

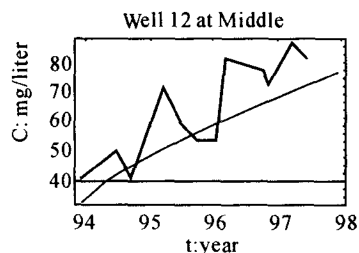


Fig.7

Tab.4 comparison of given and predicted concentrations of Calcium

Time	93.9	94.2	94.5	94.7	95.2	95.5	95.8	96.0	96.2	96.8	96.9	97.2	97.5	97.7	97.9
Given rude value:															
MW3M	41	42	44	42	49.7	36.5	19.2	22.2	62.4	60.2	49.6	63.8	97.7	85.4	61.4
MW7M	50	50	44	45	49.7	54.3	53	22.2	65.1	61.9	49.6	125	97.7	85.4	115
MW11T	39	43	50	41	41.9	44.7	43.2	48	50.7	53.3	23.5	53.2	57.3	63.6	63.8
MW12M	42	47	51	41	72.1	59.5	54.7	55.2	82.4	78.2	72.4	87.6	81.5	88.9	88.4
Predicted value:															
MW3M	31.6	35.1	38.1	40.3	44.3	46.5	48.2	50.0	50.9	54.8	55.4	57.6	59.4	60.9	62.4
MW7M	30.3	33.6	25.5	26.3	34.9	44.6	54.0	64.2	69.5	89.0	91.5	99.9	105.	109	112.
MW11T	32.0	35.8	39.1	41.6	46.1	48.6	50.7	52.8	53.9	58.6	59.3	62.2	64.4	66.3	68.4
MW12M	33.9	38.8	43.0	46.1	51.7	54.8	57.3	59.8	61.2	66.7	67.5	70.8	73.2	75.4	77.7

The thick curve is plotted according to the given data. The thin curve is plotted according to the data predicted by our model.

According to the figure above, some conclusions can be drawn as below,

(1) For near-ideal conditions, the model is suitable; for regular use, a more robust model is desired.

(2) Even though the two curves do not fit very well, they show the similar change tendency.

Sensitivity Analysis

We conduct a rudimentary sensitivity analysis, to explain the stability of our model. We vary the values of the constants α_L, α_T, θ, and V_d by 10%. Then, we can compute the corresponding changes on the values of the location and time of the pollution source, which is listed as the following (Tab.5):

Tab.5 Effects of perturbations of the parameter values

Parameter	Change of location(ft)	Change in time(year)
θ	70	0.2
α_L	<10	<0.1
α_T	10	<0.1
V_d	<10	<0.1

For the parameter α_L, α_T, and V_d, the model demonstrate a good stability. While the parameter θ has a relatively significant influence on the result of the model. Thus, it is reasonable to consider the parameter θ as a variable and repeat our grid-searching algorithm in a five-dimension space of θ, x_0, y_0, t_0 and Y. For calcium, we get the comparative results shown in Tab.6.

Tab.6 Comparison of 4-and 5-dimensional models.

Dimension	θ	X_0(ft)	y_0(ft)	t_0(year)	Y
4	0.785(constant)	7750	6040	93.75	2,100,000
5	0.84(variable)	7750	6100	93.60	2,200,000

According to the data above, the real value θ is 7 percent bigger than the estimated value. Regarding the fluctuation of the value θ, we think that there is some deflection of the direction of the groundwater flow as Fig.8.

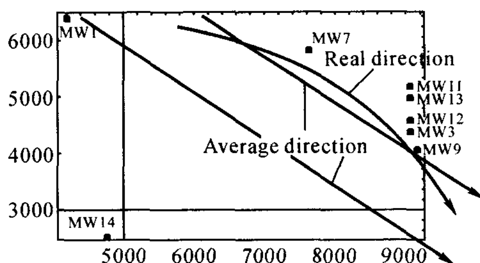


Fig. 8

Problem Two

Local Assumptions

(1) The storage tanks are located underground in the saturated zone.

(2) The direction of the groundwater flow remains the same.

(3) The saturated zone is semi-infinite.

(4) The leak process is continuous, since the primary cause of leaks in steel underground storage systems is corrosion.

Analysis

In hydrogeologic investigations of groundwater contamination incidents, monitoring wells are frequently used for sampling. A number of well configurations have been designed to meet the various sampling needs and to evaluate the most effective monitoring configurations.

Underground product leaks occur as a result of mechanical or corrosion failures in storage tank systems.

Model Design

To detect the pollutant rapidly and accurately, we develop a three-step method. First, According to the shape and size of the storage and the direction of the groundwater flow, we determine the number and location of the first group of wells. Provided that the storage is a square with a S meters side, the number of the first group of wells is defined as $N = S/20$. That

is, we drill a well every-20-meter in a line which is vertical to the direction of the groundwater, as Fig. 9 depicts:

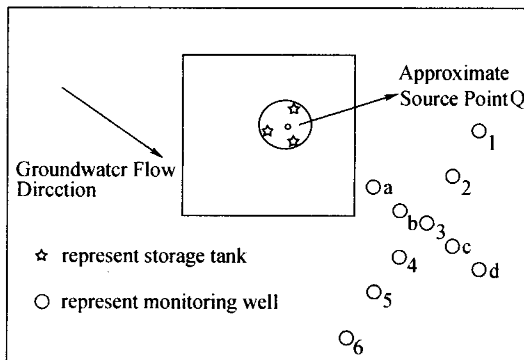


Fig. 9

In the second step, we keep documenting the data from the wells. Once there is some evidence of pollution, we start our second step. In this step, we firstly determine which well is most greatly affected by the pollutant (provided well 3). Then, from this point, we put a series of wells (perhaps five or more) along the direction of the groundwater flow. Thus, we can construct a three-dimensional formulation to calculate the fluctuation of the pollutant concentration. Different from (1), the area occupied by the storage facility may not be very large (with side less than 1000 feet). We can not use (1). Here, we employ a three-dimensional equation as follows:

$$R_d \frac{\partial C}{\partial t} + V_d \frac{\partial C}{\partial x} = D\left(\frac{\partial^2 C}{\partial x^2} + \frac{\partial^2 C}{\partial y^2} + \frac{\partial^2 C}{\partial z^2}\right) + \frac{m}{n} \tag{3}$$

Because the leaking is always a continuous process, we assume that the pollution source is applied as a step function (continuously) with the following boundary condition:

$$\begin{cases} C(x,y,z,0) = 0, (x,y,z) \neq (0,0,0), \\ m(x,y,z,t) = q \cdot C_0 \delta(x,y,z), \\ C(\mp\infty,y,z,t) = C(x,\mp\infty,z,t) = C(x,y,\mp\infty,t) = 0, t \geqslant 0. \end{cases}$$

For the case of an instantaneous point source at time $t=0$, this equation possesses an analytical solution of the form

$$C(x,y,z,t) = \frac{R_d q C_0}{8\pi n D_r} \exp \frac{V_d x}{(2D)} \cdot [\exp(\frac{V_d}{2D}) \cdot \operatorname{erfc}(\frac{r+V_d t}{2} \cdot (\frac{R_d}{D_t})^{1/2}) + \exp(-\frac{V_d}{2D}) \cdot \operatorname{erfc}(\frac{r-V_d t}{2} \cdot (R_d/D_t)^{1/2})] \quad (4)$$

Where, $r=(x^2+y^2+z^2)^{1/2}$

When $t \to \infty$, a steady-state equation can be described as following:

$$C(x,y,z) = [(R_d \cdot qC_0)/(4\pi n D r)] \cdot \exp[-V_d(r-x)/2D] \quad (5)$$

For convenience of description in the following, we employ symbol $C_m(x,y,z,t)$ to represent the right side of the (4).

For constant V_d, R_d, n, q, D, we can draw an equal-concentration plane with the concentration value $0.01C_0$ as Fig. 10.

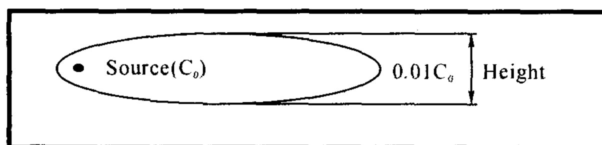


Fig. 10

Let Height be the maximum height of the equal-concentration plane.

Then, we transform the Cartesian coordinate in the same way as Problem One.

Assume the coordinate of the monitoring well is (x, y) and the thickness of the aquifer is b, we consider the concentration in the well for three situation:

(1) If $b \ll$ Height or $b \ll$ Size of the storage facility, we can transform (4) into two-dimensional equation like (1).

(2) If $b \geqslant$ Height/2, it is reasonable to consider $b=+\infty$. Thus, the problem can be simplified. We assume that the substance in the aquifer can not enter the unsaturated zone except the source point. Thus, $\partial C/\partial z|_{z=0}$

$=0$. Moreover, for every point (x, y, z) under the water-table, the concentration is doubled that which is depicted by (4) in the case of semi-infinite space.

$$C(x, y, z, t) = 2 \cdot C_m(x, y, z, t) \tag{6}$$

(3) Otherwise, for the point (x, y, z) on the upper surface of the aquifer or lower surface:

$\partial C / \partial z |_{z=0} = \partial C / \partial z |_{z=-b} = 0$

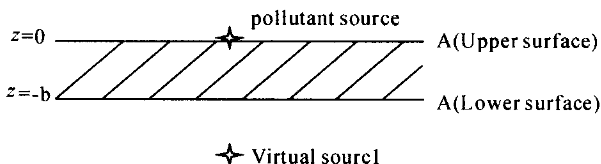


Fig. 11 Pollution source on the aquifer

Draw virtual source 1 symmetric to pollutant source with the axis being A' (lower surface). Thus, condition of $\partial C / \partial z |_{z=-b} = 0$ has been satisfied while condition of $\partial C / \partial z |_{z=0} = 0$ has not been satisfied. In the same way, we draw virtual source 2 symmetric to virtual source 1 with A (upper surface). Repeat this process, we get virtual source 3, etc. Finally, the concentration on the upper or lower surface can be considered as the result of accumulation of all the concentration value of source(including the virtual source). That is,

$$C(x, y, z, t) = 2 \cdot \sum_{0}^{\infty} C_m\left(x, y, z + 2 \cdot (-1)^{i+1} \left[\frac{i+1}{2}\right] \cdot b, t\right) \tag{7}$$

In the real model, we only need to consider the former three virtual source for the following reason:

(a) The distances from these three sources to A (A') are the most least. Their values as component affect most greatly on the value C. Moreover, other virtual sources are very far away from A (A'). Generally, the



distance between them is larger than the value Height. Therefore, we neglect these virtual sources.

(b) The pollutant discharged from the virtual sources with a long distance to the surfaces of the aquifer need a large amount of time to reach the aquifer. The longer the distance is, the more time the pollutant needs.

Finally, we transform the Equation 7 into

$$C(x,y,z,t) = 2[C_m(x,y,z,t) + C_m(x,y,z+2b,t) + C_m(x,y,z-2b,t)] \quad (8)$$

Thus, we get the final analytical solution of $C(x,y,z,t)$. Then, we use the same computer-based method as the problem A to calculate the approximate location and the time of the pollutant source.

In the last step, we draw a circle with the center of approximate source point Q and the diameter of 25 (or more) meters. In this circle, we sample some soil from the surface and analyze its chemical constituent to find the maximum. Thus, we can accurately identify the location of the pollutant source.

Monitoring Well Selection

We firstly choose those producing wells which are using now so that the water sample can represent the real chemical constituent in the aquifer.

Numerical Integration Scheme

To calculate the leakage amount, it is necessary to integrate the values of the dependent variable (C) over space. Unfortunately, the integral of (2) does not possess an analytical representation and must therefore be integrated numerically. We employ a three-dimensional integration scheme for this model. The molar mass (M) of leaked liquid is computed as

$$M = \iiint C(x,y,z,t)\mathrm{d}x\mathrm{d}y\mathrm{d}z \approx \sum_{i,j,k} C_{i,j,k}\Delta x\Delta y\Delta z \quad (9)$$

Where C_{ijk} refers to the computed concentration in the ith of x-coordination, j^{th} of y-coordination and k^{th} of z-coordination "differential" ele-

ment. We employ uniform spatial steps of $\Delta x = \Delta y = \Delta z = 1$ meter.

A Better Method for Mass Estimation

While processing the data by computer program, we minimize the variance to get a quasi-optimal solution. Meanwhile, we have estimated the m value, thus, we can compute the molar mass of leaked liquid more conveniently and efficiently as follows:

$$M = mt \tag{10}$$

Strengths and Weakness of the Model

Strengths

● The model has quite good practicality, and the given algorithm has little time complexity. For the given problem size, our C program for the grid-search algorithm runs in less than 2 minutes on a pentium-166 computer.

● The model gives good agreement of the tendency between the predicted values l and the rude data. It is fast, efficient and stable.

● As the given data are refined to simplify the computation, the accuracy will not decrease. For illustration, we list the data for Calcium as in Tab. 7.

Tab.7

Number of data	x_0(ft)	y_0(ft)	T_0
60(primitive)	7750	6060	93.70
36(after refining)	7750	6040	93.75

Weakness

● According to our assumption and computation method, if the detected area is not large enough, there is some error. As the distance between the pollution source and the monitoring well elongates, the computation accuracy increases while the measurement accuracy decreases and the response time increase. It is a contradiction difficult to solve.

● To decrease the complexity of the computation, we simplify the groundwater flow net, which may affect the accuracy of the results.

● Not taking statistical factor into our model makes the result of our model not fit the crude data exactly.

References

[1]J. H. Guswa, W. J. Lyman, A. S. Donigian, T. Y. R. Lo, E. W. Shanahan. 1984. Groundwater contaminant and emergency response guide. Park Ridge, NJ:Noyes Data Corp.

[2]Michael Barcelona, Allen Wehrhann, Joseoh F. Keely, Wayne A. Pettyjohn 1990. Contaminant of ground water: Prevention, Assessment, Restoration. Park Ridge, NJ:Noyes Data Corp.

[3] Linda M. Abriola, editor. 1989. Groundwater contamination. Wallingford, U.K.: International Association of Hydrological Sciences.

[4]C.H. Ward, W.Giger, P.L. McCarty. 1985. Ground Water Quality. New York: Wiley.

点 评

环境污染的评估是现实生活中的一个重要问题,但也是理论上较难处理的一个问题。其中的困难之一在于:这些地区往往不能进入或不允许进入,只能在周边地区收集数据,本题遇到的就是这样一种情况。题目要求根据周围抽样取得的地下水成分分析数据(系残缺的、包含错误的数据)来判断储存在某地的有害液体是否发生了渗漏,如果这种渗漏发生了,要求确定渗漏的精确位置和渗漏数量,并要求选手设计一种能发现污染源的有效检测方法,用以建立一个更为有效的数学模型,以便用于判断有害物质的渗漏是否已经发生以及发生在何时何地。这是一个关联到数学、化学、环境科学与环境工程学等多学科的跨学科课题。1999 年美国大学生数学建模竞赛首次设立多学科交叉赛题,并单独设立了 ICM 竞赛,采用本题作为赛题,较富有挑战性。

题目提供的污染物浓度数据共有上千个,在建模前必须仔细地选甄

别。本文作者显示了很强的数据分析和数据处理能力，他们未被大批数据所吓倒，未被数据文件中的残缺或重复所难住。首先，他们剔除掉了与建模关系不大的数据，例如那些浓度近似为常数、浓度变化近似随机、以及浓度显著下降的化合物数据。找出了与检查是否渗漏关系较大的化学元素，这些化学元素总共只有 4 种。接着，他们进一步考虑了地下水流的方向，排除了一些关系较小的监测井。通过上述步骤，使建模所需的数据量得以大大降低。

文章认为污染物的运动包括对流、扩散和阻滞。作者选用对流—反应扩散方程作为污染物在扩散过程中的运动方程，(这是一个二阶偏微分方程)，以此做为数学模型的基础。虽然几位参赛学生赛前均未学过偏微分方程，但他们能很快理解该方程，并将它用于研究面临的问题，体现出他们有很高的领悟力及对数学知识很强的运用能力。作者从文献中找到了点源情形下的解析解，用最小二乘法(最小变差)结合格点搜索算法确定出污染源的位置和扩散开始时间，较好地解决了问题。在求解方程时，参数的辨识非常重要。作者根据问题实际提出了一些合理的假设，从文献和互联网中找到了相关资料，并将其运用于自己的研究工作之中，从而较理想地确定出了参数的值。通过编程计算，文章认为有四种污染物泄漏，泄漏处坐标接近于(7077,6538)，(6931,5823)，(7750,6040)，(6423,7461)，前两个发生时间为 1991 年，后两个发生时间为 1993 年底。此处，通过灵敏度分析，作者们还验证了模型的稳定性。

为了快速、精确地发现污染物质(第二问的要求)，文章设计了一个三阶段方法。首先根据存储罐的形状、大小以及地下水流动的方向，确定第一组井的位置和数目(在地下水流方向的垂直线上每隔 20 米打一口井)，监测这些井的数据。一旦有污染的证据，接着就要确定哪个井受污染的影响最大(可以认为，该井位于污染传播的中轴线上)。第二步在该井的附近沿地下水流方向打一系列井作进一步监测。文章建立了一个数学模型来计算污染物浓度的波动，该模型是一个三维的二阶偏微分方程。同问题 1 一样，类似分析可以得到解析解，并用来计算污染源的近似位置和发生时间。最后一步是画一个以近似污染源点为圆心，25 米为半径的圆，在此圆中，从地表取一些土样，并分析其化学成分的变化，从而精确识

别污染源的位置。

作者在研究过程中有创见地提出了许多子问题。通过解决这些子问题,将研究逐步引向深入,并且找到了能得到有用答案的数学模型和解法。这种研究方法和写作方法使得作者能清楚地、令人信服地表述出他们的研究思路和研究结果,使论文显得层次分明,丝丝入扣,并给人以各种启发,这种做法往往正是出题人和评卷人所期望的。文章中恰到好处地使用了一些图表,使有关假设、建模、方法和结果的表述都显得更加清楚。该年的评卷人在对本文的评价中写道:作为一个周末 3 天的工作,本文给人留下了深刻的印象(quite impressive)。

本文获得了 1999 年美国大学生数学建模竞赛特等奖兼美国运筹与管理科学学会奖(INFORMS 奖),这是我国高校自 1989 年参赛以来首次获得 INFORMS 奖。

Air traffic control

李 欣 王 涛 李柏阳

Summary

Taking full consideration of the practical background of the original problem, we start our analysis by making several important assumptions. We reduce the original 3-D case to a 2-D one, since the actual airspace is usually divided into stirps of different altitudes. We base our later discussion exclusively on this assumption.

We develop a model to evaluate the probability of conflict between two airplanes. We also point out that there exists an optimal trime to resolve potential conflicts which should be a trade-off between safety and effciency.

Grasping the concept of complexity is the core of our discussion. We reason that the average number of required adjustment times per unit time suitably reflexes the "intensity" of controller workload, thus it is a fair estimation of complexity. The number of potential conflicts and the average number of required adjustment times are linked by the "right-of-way" rules. By applying the statistical method and computer simulation, we acquire a quantitative relationship between the complexity and the number of aircraft within the region. We also analyze a group of data in real case and find that the complexity varies during different intervals of day. The stability of the simulation results is also tested.

Finally, we write a summary for the FAA analyst to justify our conclu-

sions, which strives to reason that the software will definitely improve safety and reduce air traffic controller workload.

Background

In modern society, air traffic control has become of paramount concern because of the tremendous increase in the use of aircraft. However, control of air traffic is considerably difficult because it must be done for craft moving at very high speeds within a three-dimensional framework. The most important objective is to ensure that safe distances are maintained between aircraft, otherwise conflict may arise. Thus it is desirable that a software be designed to help air traffic controllers manage the increasingly complex air traffic and benefit air traffic controllers by reducing stress and workload.

By keeping controllers in the loop, the software may combine the analytical power of automation with the experience of human controllers. The software such as CTAS[1] uses real time flight plan, radar tracks, and weather information to constantly adjust to changing situations. It may increase controller situational awareness and allows the controller complete control at all times.

However, it should be kept in mind that any automation software is just a decision support tool. It assists controllers in planning the future flow of traffic in order to increase efficiency while reducing controller workload. Maintaining afe inter-aircraft separation remains the responsibility of the controller. Also the human controller must be available to catch any unresolved conflicts.

Assumptions and Justifications

Taking the practical air traffic control situation into consideration, we make several assumptions regarding the cases we deal with:

- Usually, controlled airspace is divided into flight levels[2]. A flight level (FL) is a strip of air with determinate vertical thinkness. The thickness,

however, may be neglected in our discussion. Thus we assume that during the flight each aircraft is always assigned to a given planar level and will not change altitude arbitrarily.

- There exists no interference between two aircraft of different levels. The controller only consider the potential conflicts among aircraft in the same level.
- We assume that the velocity of each aircraft is constant in magnitude during the flight. As free-flight trajectories typically will be very direct and have few turns, we assume that the direction of the velocity is not changed during normal cases unless necessary.
- Comparing to the vastness of the controlled region, the size of the aircraft can be neglected and taken as geometrical points.
- Although in real cases certain kinds of aircraft may be free from Air Traffic Control (ATC), here we only deal with those aircraft flying within the controlled region and under effective control.
- Although the actual controlled sector may vary, for simplicity here only consider the controlled region as a strip of square zones. Generally speaking, a controlled region will not be very large, usually around $100x100$ square nmi. We take a region of this size as a typical controlled region and base all of our discussions on the following coordinate, as illustrated in Fig.0:

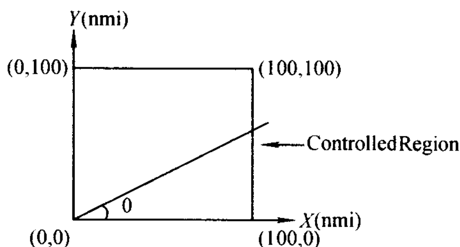


Fig.0 Our Discussion Region



Analysis of the Problem and Model Construction

Prediction of Conflict Between Two Aircraft

According to the geneal ATC regulations, above 29,000 feet, the minimum allowed horizontal separation for en-route aircraft is normally 5 nmi. If the distance between any two aircraft is predicted to fall into this spearation limit, it will be detected by the computer software and passed on to the controller. The controller makes the adjustment with the aid of computer and then instructs the pilot to make proper maneuver to avoid the conflict.

In the two-aircraft case, our primary concern is to optimize both the prediction efficiency and certainty (of course without sacrificing the safety requirement). However, the real case is that the farther in advance a maneuver is initiated, the more efficient it is likely to be, but the less certain will be exactly what maneuver is required or whether a maneuver is required at all. On the contrary, the later a maneuver is initiated, the more certain will be exactly what maneuver is required, but the less efficent and more hash the maneuver is likely to be. Thus there must exist an optimal time and a corresponding conflict probability valve value which is a trade-off between efficiency and certainty.

Usually, the approximate trajectory of any aircraft can be detected given its present velocity and location. However, due to certain unpredictable factors, such as wind influence and detection errors, the exact trajectory the aircraft follows will deviate from the predicted trajectory. It is found that the trajectory prediction errors can be modeled as monrmally distributed, with the mean value locating at zero. The unstablized along-track standard deviation tends to grow approximately linearly 0.25nmi/min. [1] For cross-track deviation, the value remains 0. 5nmi. [1] Thus the normally distributed prediction errors can be represented as ellipses in the horizontal

plane, with the major principal axis lying in the along-track direction and the minor principal axis in the cross-track direction, as is illustrated in Fig. 1.

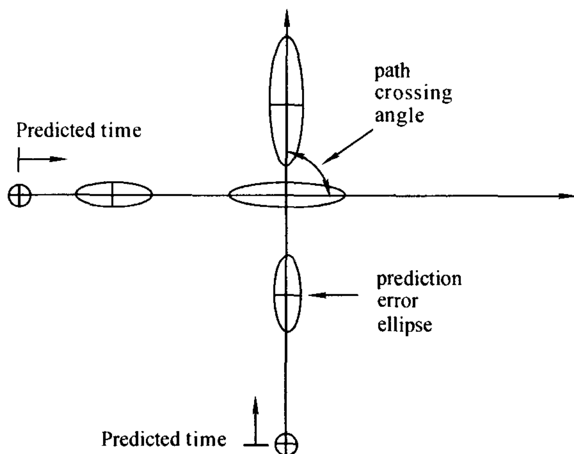


Fig. 1 The Prediction Error Ellipse

However, since, it is relatively difficult to deal with the elliptical case, or simplicity, we "extend" the ellipse along the minor principle axis and consequently just consider the error distribution region as a circle, whose radius is the length of the original major principle axis of the ellipses. However, this simplicity doesn't influence the safety requirement at all.

Given the initial location (coordinates) of two aircraft and their initial velocity (By "initial" we mean the time at which the detection occurs), we may write the expected distance between the two aircraft at any given time as follows:

$$E(l_{1,2}) = \sqrt{(x_{1,0} - x_{2,0} + vt\cos\theta_1 - vt\cos\theta_2)^2 + (y_{1,0} - y_{2,0} + vt\sin\theta_1 - vt\sin\theta_2)^2} \quad (1)$$

And the standard deviation of the distance at this time is

$$\sigma(t) = \sigma(0) + at \quad (2)$$

Where

$l_{12}(t)$ = the distance between two aircraft 1# and 2# *at t*.

$E(l_{12})$ = the expected distance between airplane 1# and 2# *at t*.

$x_{1,0}, y_{1,0}, x_{2,0}, y_{2,0}$ = the respective inital Decartesian coordinates of aircraft 1# and 2#.

V = the speed of the two aircraft, which is taken as a constant as 430nmi/hr.

θ_1, θ_2 = the respective deviation angles of the two aircraft (θ is defined as the crossing angle between the velocity orientation and positive direction of X axis), which is also independent of time.

$\sigma(t)$ = the deviation of l_{12} at time t.

$\sigma(0)$ = the inital deviation of l_{12}, taken as 0.5nmi.

a = the growing rate of the deviation of distance.

Since the deviation of the distance is the synthesis of the respective deviations of the two aircraft, it is justifiable that

$$a = 0.25\sqrt{2} = 0.35 nmi/\text{min} \quad (3)$$

Let $dE(l_{12})/dt = 0$, we get the expected miniumum distance between the two aircraft 1_{min}, and the corresponding time t_{min}.

The conflict probablility is calculated as follows;

$$P_e(t) = \int_0^{d_c} \rho(l_{1,2}(t))dl \quad (4)$$

Where d_c is the minimum horizontal conflict separation (**MHCS**), taken as 5nmi. It then can be seen that the conflict probablility is also a function of time for a specific aircraft pair.

Here ρ is the probability density of $l_{1,2}$, which is modeled as the following normal (Gaussian) distribution:

$$\rho[l_{12}(t)] = \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{ -\frac{[l_{12}(t) - E(l_{12}(t))]^2}{2\sigma^2(t)} \right\} \quad (5)$$

When $l_{1,2}$ reaches its minimum value at t_{min}, the collision probability must also be very close to its maximum.

Usually, if the minimum possible distance of two aircraft is much more than the **MIICS**, the conflict probability will be so negligible that we can be

sure to eliminate the aircraft pair from consideration. Adjustment measures are only made when the conflict probability reaches a vlave value (say, over 40%~50%).

However, it is not always desirable to make any adjustment at t_{min}, as is stated above. Although the adjustment must be made before t_{min} if the collision probablility is significant enough for the controller to take into consideration, there must exist an "optimal" time t_{opi} at which a trade-off between efficiency and certainity is achieved. It is, however, very difficult to make a very accurate evaluation of this optimal time without enough information and data. So here we can only make a rough estimation.

We assume a valve probability P_v and stipulate that an adjustment measure should be taken if the conflict probability is more than P_v. The actual choice of P_v varies in real cases. Usually P_v should be no more than 60%. This time is considered to be the optimal moment to make adjustment. Then from (4) we may solve the corresponding time at which P_v is reached. This time is considered the optimal moment to make adjustment decisions.

Multiple Aircraft Flying Through a Controlled Region

The Detection and Adjustment of Air Traffic Conflict

As has been stated above, the software detects potential conflicts and alerts the controller. Conflict prediction is based on a comparison of all aircraft pairs every given interval. After receiving this information, the controller makes proper adjustment with the aid of computer and then passes the final adjustment instruction to the pilot.

Measurement of Complexity

It is obvious that the number of potential conflicts increases as the



number of aircraft grows, while the more the number of potential conflicts is, the more the expected adjustment times there will be. Since the controller workload is directly related to the adjustment times, it is justifiable to assume the complexity is directly related to the adjustment times from an air traffic workload perspective. We therefore measure the complexity quantiatively as follows:

(1)for a particular given instant, the complexity is measured by finding the average total mumber of adjustment times at this instant;

(2)for a given interval of time period, the complexity is measured by finding the statistical average value of the adjustment times.

As is stated above, the number of required adjustment times at any given instant suitably reflexes the "intensity" of the workload, thus it is a fair estimation of complexity.

Right-of-way Rules[3]

According to the regulations set up by the Federal Aviation Agency, an aircraft in distress has the right-of-way over all other aircraft. There are three cases encountered in the right-of-way rules, as are illustrated in Fig.2:

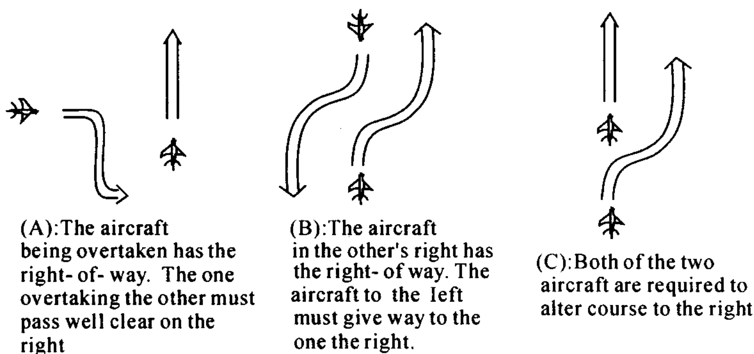


Fig.2 The Graphic Demonstration of the Right-of-way Rule

Any real case can always be classified as one of the above situations

based on the following rules. Mark the crossing angle between the directions of two aircraft as $\phi(=\theta_1-\theta_2)$:

(a) $\phi\in[45°,135°]\cup[-135°,-45°]$, t falls into the case of Fig. 2 (A), where the two aircraft fly in the orthogonal direction;

(b) $\phi\in[135°,180°]\cup[-180°,-135°]$, It falls into the case of Fig. 2(B), where the two aircraft fly in opposite directions;

(c) $\phi\in[-45°,45°]$, It falls into the case of Fig. 2 (C), where the two aircraft fly in same directions.

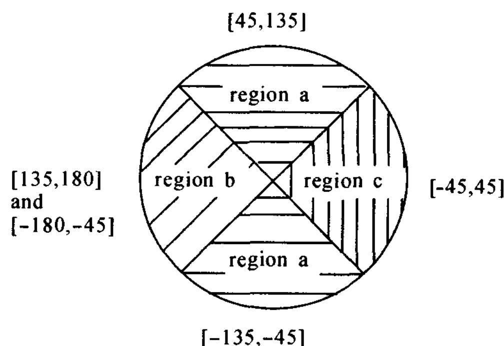


Fig. 3 The Classification of Real Cases

Normally, since the crossing angle of the two aircraft in conflict is arbitrary, it should be true that the chance a specific crossing angle occurs is evenly distributed within $[-180,180°]$. Thus the occurrence probabilities of the above three cases are 1/2, 1/4, 1/4 respectively. As is illustrated in Fig. 2, only one aircraft is needed to change route in both case (A) and (C). In case (B), however, both routes of the two aircraft are to be adjusted. Thus the statistical average adjustment times is

$$E(T_a)=(1/2*1+1/4*2+1/4*1)*E(m)=1.25*E(m) \qquad (6)$$

Here $\quad m$ = the number of potential conflicts

$E(m)$ = the statistical average number of potential conflicts.

Suppose n observations are made, where n is a suitable large number, then $E(m)$ is given by



$$E(m) = \frac{1}{n}\sum_{i=1}^{n} m_i \tag{7}$$

Where m_i is the number of potential conflicts in the ith observation.

T_a = the number of adjustment times, which depends on m.

Based on the above justifications, we apply the right-of-way rules and get the statistical average value of adjustment times for a given number of aircraft simultaneously passing through the region through computer simulation.

Suppose the number of aircraft simultaneously passing through the region is N at any given instant. We acquire the potential conflict times through simulation.

After 1000 random tests, we find that if the number of aircraft is relatively small, in most cases the conflict times is tiny. As the number of aircraft increases, chances of conflicts also grow.

Simulation A: Potential Conflicts Tiems At Any Given Instant

We arrange our simulation process into certain phases. In each phase, the number of aircraft is held as constant. Here we begin with $N = 2$.

During each phase of simulation, we generate N aircraft randomly within the region ($100x$ 100nmi) along with their locations every time. The speed of velocity is always maintained at 430nmi/hr. It is also ensured that the distance between any aircraft pair is larger than the **MHCS** at the time of generation. We then predict whether conflict of any aircraft pair will arise within our given period of time (taken as 10min) and record the total number of potential conflicts m. Such observation is made 1000 times in each phase, and the occurrence frequency of m conflicts $F(m)$ ($m = 0, 1, 2, \cdots$) is counted.

Tab.1 The Results of Simulation A

N \ $F(m)$	m								
	2	3	4	5	6	7	8	9	10
0	921	755	576	444	261	152	107	45	17
1	79	226	351	394	423	373	264	178	95
2	0	15	53	106	200	252	275	256	213
3	0	4	14	36	83	13	166	199	193
4	0	0	5	15	23	62	102	139	193
5	0	0	3	4	5	14	46	97	108
6	0	0	1	1	1	10	20	36	84
7	0	0	0	0	4	1	14	20	48
8	0	0	0	0	0	1	5	19	27
9	0	0	0	0	0	0	1	6	11
10	0	0	0	0	0	0	0	2	5
11	0	0	0	0	0	0	0	1	2
≥12	0	0	0	0	0	0	0	0	0
$E(m)$	0.079	0.268	0.540	0.802	1.223	1.669	2.217	2.921	3.639
$E(T_a(N))$	0.099	0.335	0.675	1.002	1.529	2.086	2.771	3.651	5.114

In the next phase we increase N by 1 and repeat the same process until N reaches 10, which we regard as the upper limit of the number of aircraft that can be flying through the region simultaneously. The value of $E(m)$ and $E(T_a)$ are also calculated using (6) and (7). Following are the results of our simulation:

From the above data we may find that the number of potential conflicts, thus the average number of adjustment tiems increases as the number of aircraft flying through the region grows. It is also noted that the possiblity that more than one pair conflict occurs also grows with the increasement of N.



Aircraft passing Through The Region During A Given Interval of Time

The case that aircraft pass through the region during a given interval of time is far more complex and uncertain than the former situation. The main difficulty lies in:

i) The number of aircraft flying within the region is constantly changing with time due to arrival and departure flux;

ii) As the time passes by, new conflicts may arise in spite of the already solved ones. Thus continuous detection and adjustment become necesary.

Due to these analysis, we write the number of aircraft within the region at instant t as the function of time, which is $N(t)$. Suppose the given time length is T. Since complexity is measured by the adjustment times per unit time, it is necessary that the average number of aircraft flying within the region per unit time during T be calculated, as is given below:

$$E(N) = \frac{1}{T}\int_0^T N(t)\mathrm{d}t \tag{8}$$

We divide T into m equal intervals designated as Δt. Each of these intervals is so short that we may reasonably assume that the number of aircraft flying within the scope of the region is a constant during each Δt, designated as $N_1, N_2, N_3, \cdots, N_i, \cdots, N_m$. All the N_i aircraft spend all of their corresponding Δt exclusively in the region. Thus the total time all the aircraft spend in the region during T may be approximated as:

$$t_{total} = \sum_{i=1}^{m} N_i \Delta t \tag{9}$$

let $\Delta t \to 0$, we get the following result:

$$t_{total} = \int_0^T N(t)\mathrm{d}t \tag{10}$$

Thus we can acquire $E(N)$ through t_{total}:

$$E(N) = \frac{t_{total}}{T} \tag{11}$$



Let N_{total} be the total number of aircraft that have ever been appeared in the region and t_i be the period that the ith aircraft spends in the region during the interval T. ($i=1,2,\cdots, N_{total}$). It is obvious that

$$t_{total} = \sum_{i=1}^{N_{total}} t_i = N_{totoal} * T_s \tag{12}$$

Where T_s is the average time that an aircraft spends in the region during T:

$$T_s = \frac{1}{N_{total}} \sum_{i=1}^{N_{total}} t_i \tag{13}$$

In order to estimate, we introduce T_p as the average time that it takes an aircraft to pass through the region. It is obvious that $T_s = F(T_p, T)$. The actual form of this function, although remaining unknown, must satisfy the following intuitive conditions:

T_c can not exceed the total interval T. Actually, if T is small enough, we can justifiably assume that the total number of aircraft within the region is a constant, thus $T_s = T$. On the other hand, if T is very large, we may think that all aircraft can enter and leave the region within T, thus $T_s = T_p$. So:

$$\lim_{T\to 0} Ts = T \tag{14}$$

$$\lim_{T\to \infty} Ts = T_p \tag{15}$$

Within any given interval T, the longer the time for an aircraft to pass through the region T_p is, the more likely that an aircraft is likely to spend more time within the region, which means that T_s will be greater. On the other hand, if T_p is fixed, then T_s will increase with the length of interval T.

Based on the above discussion, we come out a simplest model that can satisfy the above constraints:

$$T_s = \frac{T * T_p}{T + T_p} \tag{16}$$

Although it is a very rough model, one may find that it is a fairly well approximate reflection of the trend of T_s as T and T_p varies.

In our discussion, we fix speed of the aircraft, thus T_p is a constant whose value we can acquire by statistical simulation (See Appendix).

From (16),(12)and (11) we can write out the formula of E(N):

$$E(N) = \frac{T_p}{T + T_p} * N_{total} \tag{17}$$

From $E(N)$ we can get the average number of adjustment times $E(T_a(N))$ by using the results given by Tab. 1. However, the calculated $E(N)$ may not necessarily be an integer. So we have to apply the method of polynomial interpolation to get the quantitative relationship between the average adjustment ties $E(T_a(N))$ and the aircraft number N. Thus we can get the corresponding $E(T_a(N))$ at $E(N)$ directly from the fitted curve in Fig.4.

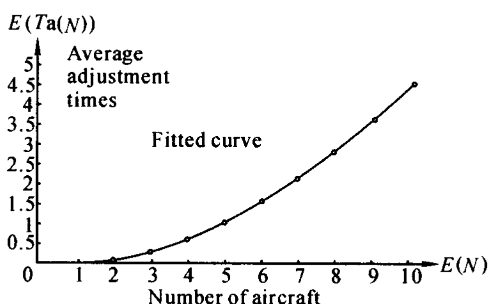


Fig.4 The Fitted Curve of the relationship between $E(N)$ and $E(T_a(N))$

The fitted formula is given below:

$$E(T_a(N)) = f(E(N)) = 0.503 * E(N)^2 - 0.0531 * E(N) + 0.0264 \tag{18}$$

Each time we give an N_{total}, from (17) and (18) we can get a quantitative relationship between average adjustment times and the interval T. This is illustrated in Fig.5:

Aircraft Passing Through The Region During A Particular Time Of Day

In reality, the number of aircraft passing through the region varies during different time of the day.

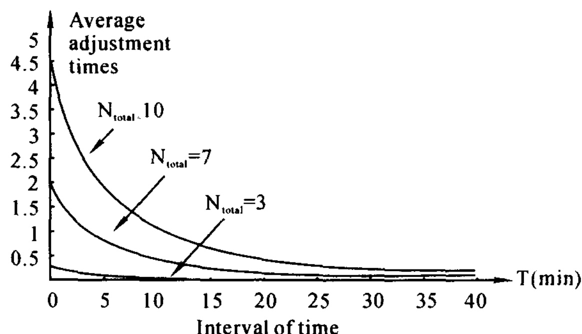


Fig. 5 the dependence of the average adjustment times on the length of interval time, given $N_{total}=3,7,10$ respectively.

We analyze the data acquired from a control tower around the International Airport is Sydney during a week. We divide a day into 12 equal intervals, each interval lasting 2 hours. We then count up the total number of aircraft passing through the region during each interval and calculate the corresponding average adjustment time by applying (18) and (17). It is found that the flux of aircraft as well as the average adjustment times per interval varies significantly during different intervals of the day. Fig. 6 shows our statistical results:

It is obvious from the above graph that during a particular day, the number of aircraft within a region, thus the value of average adjustment times is the greatest at round 2:00 - 4:00 PM. During night, however, the value of average adjustment times is very tiny. We therefore reasonably conclude that the workload complexity of controllers is the most significant at noon while the controller is more relaxed at night.

Stability Test

In solving the problem, we put the limitation that the speed is a given constant. (430nmi/hr) and get our simulation results. We now vary this situation by generating the of velocity of each aircraft randomly within the range of [390,470](nmi/hr) and run tests by repeating the same process in

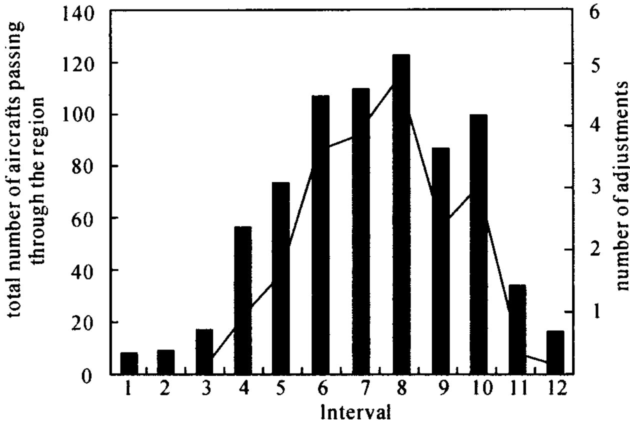


Fig.6 the number of aircraft passing through the region (represented by the bar graph) and the numbr of adjustment times (represented by the broken-line graph) during different intervals of a particular day. Each interval contains 2 hours.

simulation A. Tab.2 gives the results of the tests.

Tab.2 Results of the case in simulation A with varying speeds

$E(T_a(N)$ \ V(nmi/hr)	m								
	2	3	4	5	6	7	8	9	10
430	0.099	0.335	0.675	1.002	1.529	2.086	2.711	3.651	5.114
390～470	0.103	0.319	0.629	1.033	1.524	2.143	2.981	3.731	4.911
Relative Deviation(%)	+4.04	-4.78	-6.81	+3.09	-0.33	+2.73	+7.58	+2.19	-3.96

From the above table we find that the results we gain in simulation A is fairly stable with the speeds varying in the given range. Changes in speed have very little effect on the average adjustment times, since the relative deviation is always fluctuation around zero with tiny peaks.

Conclusions

● We find that the probability of conflict between any aircraft pairs varies with time. Taking both safety and efficiency into consideration, we believe that there exists an optimal time to resolve the conflict, at which the probability value reaches a valve value.

● We measure the complexity by finding the average number of adjustment times. Through statistical ananlysis and computer simualation, we find the relationship between the average number of adjustment times and the number of aircraft flying through the region both for one instant and for a given interval of time.

● In a day, the average adjustment times vary with different intervals. By analyzing the data available from a particular control tower, we find the distribution of the average adjustment times during different intervals of the day.

● We think that the presence of the software tools should considerably reduce the complexity (measured by the average adjustment times). Since without the aid of the software, the controller has to make subjective judgement of whether to make adjustment or not, which may result in great errors. Since the primary concern of the controller is to ensure the safety, he has to make many unnecessary adjustments. The high accuracy of the prediction of the software, however, will significantly reduce the unnecesary adjustment tiems. Thus the presence of the software will surely reduct the workload and complexity.



Strengths and Weakness of the Model

Strengths

● We take full consideration of the practical background of the original problem and make several reasonable assumptions by neglecting minor factors.

● We simplify the original 3-D problem into a 2-D case, which both accords with the aviation routine and successfully reduces the difficulty of the problem.

● We realize that the software is just a decision support tool intending to help reduce the workload of controllers. It can not fully replace the role of the controller, which also accords with the real case.

● In analyzing the prediction process, we take both safety and efficiency into consideration and give a reasonable estimatation of the proper prediction time.

● We give a reasonable criterion of measuring the complexity and get the quantitative relationship between aircraft number and complexity through computer simulation and statistical analysis. Also the results have a fair stability.

● We reasonably anzlyze the way in which the actual software works and our simulation algorithm is fast enough to gain sufficient data for analysis.

● Our original data are directly acquired from the real ATC parameters and are highly reliable. Thus our model has a considerable practical value.

Weakness

● We neglect some factors that may affect the air traffic control, such as the influence of wind and the deviation of flight trajectories due to random factors.

● In real cases, the normally distributed prediction errors should be represented by ellipses in the plane. While we simplify it to a circular case in our

discussion, which will unavoidably introduce certain deviations.

● When seeking the quantitative relationship between the average time an aircraft spends in the region during the interval T and the value of T, We put forward a rough formula which is not very accurate and is subject to further revision.

Appendix A: The Acquirement of The Average Time That An Aircraft Passes Through The Given Region

After gaining results of 1000 such tests, we get the average value of the passing time, which we take as T_p in our further discussion. Given the speed as 430nmi/hr, the result is 11.28min.

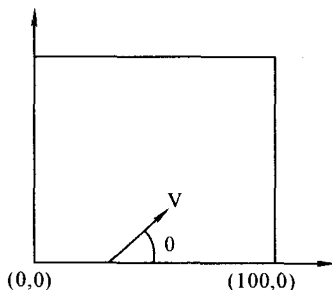


Fig.7 Simulation of the Passing Time

Reference

[1] Conflict Probability Estimation For Free Flight. Russell A. Paielli and Heinz Erzbeger. NASA Ames Research Center, Moett Field.

[2] Field Test Evaluation Of the CTAS Conflict Prediction And Trial Planing Capability. B. David McNally and Ralph E Bach, NASA Ames Research Center, William Chan, San Jose University.

[3] Private Pilot Manual. Jeppesen Sanderson, Sanderson Training Systems.

[4] Computer Simulation and Monta Carlo Methods. Fang Zai Gen, Beijing Industry College Press.

The data are taken from the CTAS website: http://www.ctas.arc.nasa.gov



Summary

(To be presented to the FAA Administrator)

The new air traffic control software is designed to improve safety and reduce air traffic controller workload. After a detailed analysis of the air traffic conflict prediction and adjustment process, we come to the conclusion that the software will significantly reduce the workload of air traffic controllers as well as improving the prediction accuracy.

A controller usually handles a particular airspace sector and takes up the responsibility to eliminate the potential conflicts within it by ensuring a suitable separation between the aircraft pair. Traditionally, the role of the controller is to both predict the potential conflicts and make proper adjustments to eliminate them exclusively from his personal experiences. Thus inaccuracy and mistakes may unavoidably occur. Since the primary concern of the controller is to ensure the safety, he has to make many unnecessary adjustments. Also chances will be likely that he will fails to predict some potential conflicts, which reduces the reliability of prediction.

The high accuracy of the prediction of the software, however, will significantly reduce the unnecessary adjustment times as well as improve the air traffic safety. It can calculate the future trajectories of aircraft from their current positions and velocities with minor errors. the software then predicts the conflict time with tiny errors. Bases on this information, the controller is able to make proper prevention adjustment which is subject to the judgement of the software. The software should work in a real-time way so as to provide up-to-date information to the controller and judge whether the adjustment measures are right or not. By keeping controllers in the loop, the software may combine the analytical power of automation with the experience of human controllers. The software may use methods such as real time flight plan, radar tracks, and weather information to constantly adjust to changing situations. It will incerease controller situational awareness and allow the controller complete control at all times.

For further analysis, we develop a model to study the factors influencing the workload of controllers and how the software reduces the controller's workload. The workload during and interval is directly related to the number of adjustment times made per unit time. We find that the main factor that influences the controller's workload is the number of aircraft passing through the controlled sector during this interval. We also deduce an approximate quantitative relationship between the number of aircraft passing through the controlled sector during any given interval and the corresponding adjustment times. We find that the adjustment times increase sharply as the number of aircraft grows. Thus the greater the number of aircraft is, the greater the extent to which the softwre reduces the workload of controllers.

It is also discovered that the flux of aircraft of the controlled region varies with different time of day. As is noticed, the value of the software become especially obvious during these rushes.

In conclusion, the software is a fairly well assistant of the air traffic controller. It can enhance the air traffic safety as well as reduce controller workload. It is therefore worthy to be put to practical use in the air traffic system.

点 评

本文作者根据飞行管理的实际背景,将其简化为一个两维正方形区域内的航线调控问题,从而降低了问题的难度与计算复杂性。他们首先导出了一个估算两架飞机碰撞概率的计算公式,并用随机生成的数据加以检验。作者运用期望值方法研究了飞机在一个给定的时间段内通过空中一个指定区域时的情况,得出了飞机的平均调整时间数,并用取自悉尼国际机场一周内统计的实际数据对结果作了检验,并就稳定性问题作了一定的研究。

本文对飞行管理问题开展了较具一般性的研究,从平均值(期望值)



角度分析问题,得出了一些能兼顾安全和效益的办法。从文章中可以看出,作者并不轻易相信自己通过理论推导得出的结果,他们能用大批随机产生数据和来自真实环境的实际数据反复地考验自己的研究结果,并进而开展稳定性分析,这种务实的研究作风是值得提倡的。

本文获得了2000年美国大学生数学建模竞赛一等奖。

The optimal assignment of channels

陈 驷 吴以均 来 翔

Summary

To find the span, the smallest one of the largest channels in all feasible assignments that complies the constraints, we model the assignment of radio channel.

We develop a backtracking algorithm, simple and efficient, to find the span of a finite area. In addition it can be used to get the lower bound of the span of an infinite area.

Our approach is to get some specific results using the backtracking algorithm at first, and then to study these results and deduce some rules from them, at last to prove them theoretically.

The backtracking algorithm works very well to the area described in the problem paper, we call it area A, finds the span is 9. After analyzing the assignments given out by the algorithm, we find it's very regular. There is an element region composed of 9 hexagon cells, with repeatedly spreading this element region we can get the assignment for the whole area.

If the area spreads infinitely, we find that the element region got from the assignment for area A still works, so the span is still 9 and the spreading process of this element region gives the method to get the corresponding assignment. It is very easy to implement.



If the channels assigned to adjacent transmitters differ by at least other numbers, denoted as k, instead of 2, we still want to find an element region. And virtually we got it, and corresponding to this element region, the upper bound of span is $2k+7$. We proved that when k is greater than 6, $2k+7$ is also the lower bound. To the other values of k, we can use the backtracking algorithm to find the lower bound of the span. If this lower bound is equal to the upper bound, then we get the span. Finally we get the result:

$$f(k)=\begin{cases}7 & k=1\\ 3k+3 & 2\leqslant k\leqslant 3,\\ 2k+7 & k\geqslant 4\end{cases}$$

where $f(k)$ denotes the span dependant on different k. To each value of k, there is a simple rule to design the assignment that attains the span.

We have completely resolved the problem when there's no more than two level interference. However under actual situation the problem is more complicated. We furthered our research on some factors that we think important or common and improved our model.

Introduction

Our goal is to assign channels to hexagon cells that cover a finite or infinite area so that we can save channels to the greatest extent while still complying with the constraints below.

- A channel can be assigned to several cells, but the distance between those cells should be no smaller than a specific value.
- Channels assigned to adjacent cells must differ by a specific value.

The specific case given assumes that no two cells within the distance $4s$ of each other can be assigned the same channel and cells within the distance $2s$ of each other mustn't be given the same or adjacent channels, where s is the length of a side of one of the hexagons.

We first developed a backtracking algorithm to find the answer for

some limited cases. And after studying the results of our program, we abstracted the rules of the problem and solved the specific problem.

Then we extended our model to solve a more generalized question. Channels within the distance $2s$ must differ by at least some given integer k. We presented a function about k, which can also accurately find the span, and give out a method to implement the assignment.

We furthered our approach to an even more generalized problem that there are several levels of interference. We solved this problem to an acceptable extent, giving out the rules when the number of levels is 3.

What is more, we also discussed some other aspects of the problem, such as irregular transmitter placement.

Background and Assumptions

In a cellular communication network, channels are a precious resource and we should carefully plan its usage. Naturally we would like to use as few channels as possible, but the difficulty of interference adds constraints to our plan. The two most serious forms of interference are adjacent channel interference and co-channel interference.

With frequency reuse, the same channel will be used by several cells. These are called co-channel cells and obviously they can cause interferences if placed too near. So, co-channel cells should be placed far away enough to avoid co-channel interference.

Adjacent channel interference comes from the fact that there is always some part of a channel's spectrum interfering with the next channel. Thus, these adjacent channels should be arranged as far as possible in cells, to reduce the probability of interference.

When allotting channels to the cells, our goal is to minimize the span of the assignment while avoiding the two kinds of interference.

Our model design is based on these assumptions:

- The area is which is covered by the cells is planar and continues.



- The transmitters are located at the center of each cell.
- The area has a honeycomb-style shape shown as figure 1 and all the cells are hexagons and the same in size.
- Channels are regularly spaced so that they can be represented by integers 1,2,3, etc.

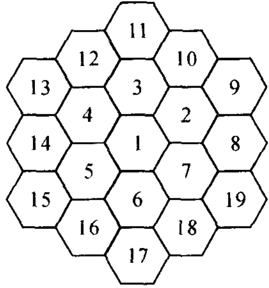


Fig. 1

Symbol Table

Below is the symbols we have used in our model:

Symbol	Explanation
$d_{i,j}$	The distance between the center of cell i and cell j.
l	The level of interference (In requirement *A*, *B* and *C*, it is 2).
k_j	The distance that two cells must differ if there exists j level interference between them (In requirement *A* and *B*, for instance, $k_1 = 2$ and $k_2 = 3$.) If we do not specify the subscript, k means k_1. So we can also say that $k = 2$.
span	The minimum, over all assignments satisfying the constraints, of the largest channel used at any location.
t	The largest channel of one feasible assignment
s	The length of a side of one of the hexagons.
N	The number of cells in a finite area.

Model Design

A straightforward solution: Backtracking

The easiest method to come up with is to check every feasible assignment and find the smallest spectrum width, in other words, the span. This method is feasible if the area is not very large. And it is actually feasible to the case shown in Figure 2, because its size is not too large.

After assuming the span to be one value, this method works like this:

we select a cell one by one, according to some order, and assign it a channel, smaller than the assumed span, without causing interference. If this channel cannot be found for one cell, we go back to the previous cell and assign it next feasible channel, and continue to next cell, and just repeat this process until we assign every cell a channel and do no cause interference. That's to say, we find a feasible assignment, then the span is no larger than the value we assumed in advance; or we check all assignments possible and get the conclusion that span of this area is larger than the value.

By increasing the assumed value, we can always get an assignment at last. The smallest value is just the span.

Preparation

To locate each cell in this area, we establish a coordinate system as shown in figure 2. We use (x, y) to denote the coordinates of cell P, x is the distance from axis Y to the center of cell P in the direction of axis X, and y is the distance from axis X in the direction of axis Y. $(x=3, y=2)$

The distance between two centers is given by the formula:

$$d_{p,q}=\sqrt{(x_p-x_q)^2+(y_p-y_q)^2+(x_p-x_q)(y_p-y_q)}.$$

This formula is used when checking whether one transmitter is within the interference area of another transmitter. The proof is shown in appendix A.

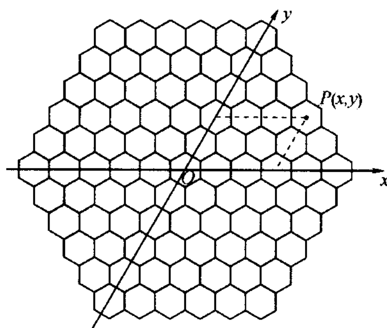


Fig. 2

We define an order between cells in the area: if $X_p < X_q$, P is before Q; or if $X_p = X_q$, their order is determined by y-coordinate, the smaller one is before the larger one. According to this order, we can assign a unique serial number to each cell, from 1 to the number of cells in the area, denoted with N and to the

area A it is 91. We can use this serial number, as well as coordinate, to identify each cell. We use j to denote this serial number.

Algorithm Implementation

Now we can implement the algorithm. Let t be the largest channel of one feasible assignment, we test the t from one enough small value, extremely 1. Specially to area A, since we can simply prove that it must be at least 9, (the proof is given in Appendix A, Deduction 2) we start our searching for span from 9, that's to say, let t increase from 9. For each t we try to find an assignment, because this is a finite area, finally we can always get a t for which there is an assignment. This t is just the span related to this area. The definite description is given below. It finds the span and gives all assignments in which the largest channel is span. The program is in Appendix C.

(1) $t = 9$, set all channel(j) to 0 ($1 \leqslant j \leqslant 91$)

(2) $j = 0$

(3) $j = j + 1$

(4) if $j = 92$, then set *span* to t, output all channels(j) ($1 \leqslant j \leqslant 91$), go to step (7)

(5) if $j > 0$, go to step (7)

(6) if *span* $\neq 0$ go to end, else $t = t + 1$, go to step 2

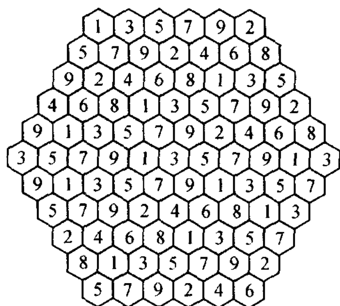


Fig.3

(7) if channel(j) = t, channel(j) = 0, $j = j - 1$, go to step (5) else channel(j) = channel(j) + 1

(8) check whether there are interference between cell j and other all cells before j, if there are not interference, go to step (3), else go to step (7)

We get the result that t is just 9. Fig.3 shows one of the correspond-

ing assignments.

Extend to infinite

So far we have already seen that it is enough to cover the finite area with a span of 9, now we are going to extend our method to infinite area.

Can an assignment with a span of 9 fulfill our task to cover an infinitely large area?

From the result of our backtracking algorithm, we find an interesting assignment, as illustrated in Figure 4.

This 9-cell structure can be duplicated and patch up like mosaics to cover an infinitely large area while still complying the constraints set by the problem, please see Figure 5.

Now we have constructed an arrangement scheme, we can say that 9 is the upper bound of the span, and since we have proved that if the span is 8, we cannot cover an infinitely large area, we actually have find the correct and accurate answer: the span is 9.

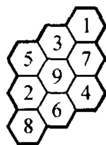


Fig.4

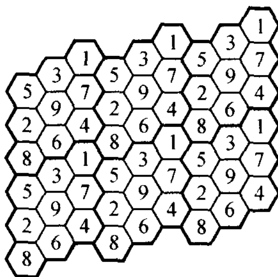


Fig.5

What backtracking can do now

If the area can spread infinitely, we cannot use the backtracking algorithm directly to find its span, but it is still useful. As lemma 1 says, if area A can cover area B, span(A) must be greater than span(B), the span of an infinite area is larger than the span of any finite area. So we can use the

backtracking algorithm to find out the span of some finite area to get the lower bound of the span of the infinite area. And this strategy is useful when considering the situation that channels for transmitters within distance $2s$ differ by at least an integer k instead of 2.

More general circumstances

Repeat requirements of the above section, except assume now more generally that channels for transmitters within distance $2s$ differ by at least some given integer k (define as Constraint 1), while those at distance at most $4s$ must still differ by at least one (define as Constraint 2).

Efficient strategies for designing assignment

Using the Backtracking for area A while k is form 1 to 8, we can get the span of area A and the assignments that attain the span for each k. These results are shown in Fig. 6 and Tab. 1.

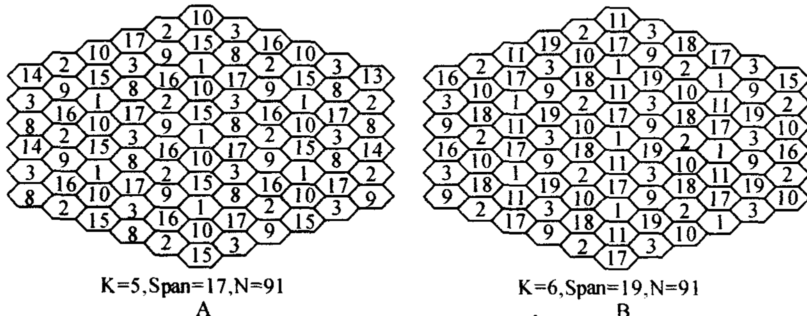


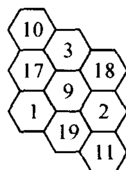
Fig. 6 Partial results from the Backtracking

Tab. 1 Results from the Backtracking

K	1	2	3	4	5	6	7	8
span	7	9	12	15	17	19	21	23

From Fig. 6, we can see some rules in these graphs. We find some element regions and by using them we can fill the entire area, satisfying the constraints. For example, one of the element regions can be drawn as the

following Fig. 7.



(A part of the graph B in Fig. 6)
An example for the element region
Fig. 7

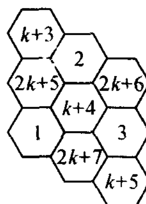


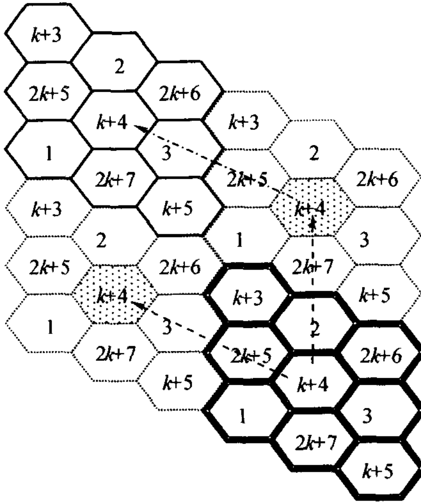
Fig. 8 Nine-cell structure

And Tab. 1 shows the *span* is $2k+7$ from $k=4$ to 8. So after studying the results of the *Backtracking*, We find the rule of the element region. It's one form of the element region shown as the Fig. 8 (Nine-cell structure). This element region embodies the features of cell cluster, which is an important concept of cellular communication[1]. Thus the Nine-cell structure owns the merits of cell cluster. The cell structure can be duplicated and patch up like mosaics to cover an infinitely large area while avoiding interferences. Also the distances between adjacent co-channels are always the same[1].

How to spread

One successful approach to solve the problem is shown as Fig. 9, like the spread of cell cluster. It can be used to spread the Nine-cell structure out. Each cell in the element region, called as the initial cell, can be spread out in six directions. In the first spread, walk over three cells in the direction perpendicular to a side of the initial cell, to the center of the third cell, the destination. This destination cell should use the same channel as the initial cell does. So each initial cell will get six copies in six directions. As the result, the element region gets six copies in six directions. Then in the second spread, each of the six copies can also get six copies if ignoring the overlaps. Going this way, basing on the characters of cell cluster, we can fill the area with the element regions (Nine-cell structure), complying the

constraints.



The broken line arrowhead shows the first spread, and the dashdotted arrowhead shows the second spread.

Fig. 9 the Nine-cell structure method

Using this method, we'll solve the problem with the largest channel $= 2k+7$. So the *span* is no larger than $2k+7$. In other words, we get an upper bound of the *span*.

Prove the function f(k)

Tab. 1 shows the *span* is just $2k+7$ where $k \geqslant 4$. So $f(k)$ maybe is $2k+7$, $(k \geqslant 4)$ and virtually it is. We can give the proof for the function f.

The lower bound of the span of infinite area

Let's firstly discuss $f(k)$ when $k > 6$.

Discussion for $k > 6$

Divide the interval of the frequency spectrum to three sects. These sects are $[1, k]$, $[k+1, 2k]$, $[2k+1, 3k]$. So there are three types of channels. Every channel in sect 1 is named as the Small channel, abbreviated as S. And each channel in sect 2 is named as the Middle channel, abbreviated as M. The Large channel is a channel in sect 3, abbreviated as L.

Using this division, we can get two simple results.

[General Information]
书名=数学建模竞赛：浙江大学学生获奖论文点评 1999-2004 普通高等教育“十五”国家级规划教材配套用书
作者=杨启帆 何勇 谈之奕编
页数=597
SS号=11484953
DX号=
出版日期=2005年07月第1版
出版社=浙江大学出版社